OF MICHIGAN
DEC 18 1961
BUSINESS ADMINISTRATION
LIBRARY

# Journal of Advertising Research

Vol. 1, No. 6 DECEMBER 1961

## COPY TESTING ISSUE

**Sales Effects of Two Campaign Themes**
PETER L. HENDERSON, JAMES F. HIND AND SIDNEY E. BROWN 2

**A Sensitive Measure of Ad Exposure**
GORDON M. KESWICK AND LAWRENCE G. COREY 12

**Measuring Advertisement Recognition** D. MORGAN NEU 17

**Recognition of Ad Elements: Recall or Projection?**
ERIC MARDER AND MORT DAVID 23

**A Pre-Test of Four Australian Paint Ads** RONALD VICKERS 26

**Three Ways to Improve Ad Pre-Tests**
BILL STEVENS AND JOEL AXELROD 33

**FEDERAL STATISTICS IN ADVERTISING** 37
**ON METHODS** 40
**RESEARCH IN REVIEW** 47
**PUBLICATIONS RECEIVED** 54
**EDITORIAL** 56
**INDEX TO VOLUME 1** 57

Published by the Advertising Research Foundation

# The Journal of Advertising Research

is published quarterly by the Advertising Research Foundation and sent free to persons in its member firms. Others may subscribe at $5 per year or buy single copies at $1.25 each. Excepted are persons in companies eligible for regular ARF membership (U.S. advertisers, agencies and media), to whom the single-copy price remains $10.

solicits original papers. Reports of findings are favored over theoretical discussion. Manuscripts should be submitted in triplicate, double-spaced, with references, tables and figures on separate pages. Authors receive 50 free offprints. Letters of comment and criticism are also invited.

is intended for practitioners and users of advertising research. Limits on readers' time and journal space require that papers be as short as clarity permits.

is an open forum. Publication in it implies no endorsement of the writer's purpose, methods or views by the Advertising Research Foundation, its board of directors, or any of its committees.

is edited by the ARF Technical Staff:

Charles K. Ramond ..... Technical Director
Ingrid C. Kildegaard ..... Research Statistician
Gwyn Collins ..... Research Mathematician
Thornton C. Lockwood ..... Research Associate
Naomi Boretz ..... Librarian
Joey Harris ..... Secretary

© 1961 by the Advertising Research Foundation, Inc.
3 East 54th Street, New York 22, N. Y.

# Contents

One aspect of advertising, like that of business generally, is strictly a matter of managing money. As a businessman the advertiser must decide how much to invest, where, when and how. But as a communicator, his most characteristic decisions are rarely as quantifiable. Here he must choose among different things, not just among different degrees of the same thing; among different approaches as well as among different emphases on the same approaches.

In this issue we focus attention on these qualitative choices which confront the advertiser-as-communicator. Each article reports or discourses on a way of selecting from many the one most effective advertisement. All six are concerned with the translation of this qualitative problem into measurable, quantitative terms.


– Henderson, Hind and Brown report a copy test long thought difficult if not impossible: the comparison of two promotional themes in terms of the sales they caused. p. 2

– Keswick and Corey describe a new measure of advertising exposure which gives an ad a higher score the more it has been exposed. p. 12

– Neu presents the case for a traditional way of measuring ad recognition, the Starch noted score, and some new evidence for its usefulness. p. 17

– Marder and David report a limitation of the recognition technique; they find that it does not adequately measure exposure to the parts of an advertisement. p. 23

– Vickers asked Australians to write their own copy for ad illustrations; his analysis of it confirms several copy testing maxims and turns up some surprises. p. 26

– Stevens and Axelrod call for a three-part covenant between user and doer of research which should improve their performance as much separately as together. p. 33


The sections written by ARF staff:


– In *Federal Statistics in Advertising,* Miss Kildegaard charts five sources of data on consumer income and shows the particular value of each. p. 37

– Collins, *On Methods,* having previously discussed correlation, regression and factor analsis, describes a concept which in a sense underlies them all, analysis of variance. p. 40

– In *Research in Review,* Collins, Lockwood and Shirey examine at length three books and five articles, as well as 13 other publications in brief. p. 47

– *Publications Received* since the last issue are listed by Mrs. Boretz, who also tells how our members use the ARF Library. p. 54

– The editorial answers some questions raised by applying computers to media selection and asks: who will do the basic research from which such applications are derived? p. 56

# Sales Effects of Two Campaign Themes

PETER L. HENDERSON, JAMES F. HIND and SIDNEY E. BROWN

U.S. Department of Agriculture

**Advertising is only one of many factors influencing purchase. Here is an experiment which took the most significant factors into account and permitted a sales comparison of the themes employed.**

SOUND PLANNING AND CONTROL of advertising requires improved measurement of advertising effectiveness.

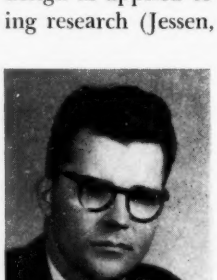
**PETER L. HENDERSON heads the Development Analysis Section, Market Development Branch, Marketing Economics Division, U. S. Department of Agriculture. He has served on the staffs of the University of Georgia, Cornell University, and Virginia Polytechnic Institute, where he taught and conducted research in marketing. He has authored and co-authored various research reports on the merchandising and promotion of fruit. He received a Ph.D. from Cornell in 1952; for his dissertation he adapted the double change-over experimental design for use in marketing research. This application, reported as "Special Promotional Programs for Apples —Their Effects on Sales of Apples and Other Fruit" by the USDA in January 1961, forms the basis for this article.**

**JAMES F. HIND received a B.S. in 1954 and an M.S. in 1958 from the University of Tennessee where he studied marketing and applied statistics. An agricultural economist in the USDA's Market Development Branch, he has used experimental designs in advertising research to evaluate the sales effectiveness of the promotion of agricultural products. He belongs to the American Marketing Association and the American Statistical Association.**

The purpose of this study was to evaluate the sales effectiveness of a specific promotional campaign. While the research techniques and statistical analyses were developed by physical scientists, the adaptation of these methods to advertising research has resulted in improved measurement of advertising effectiveness.

The basic experimental design and analysis was employed in the biological sciences as early as 1941 (Cochran et al., 1941). It was later adapted and applied to problems in market research by Henderson (1952) as cited by Federer (1955) and in *Wood Chips* (1959).

A recent article published in this *Journal* discussed the theory and concept of an experimental design as applied to a specific problem in advertising research (Jessen, 1961). A mathematical model



**SIDNEY E. BROWN has a B.A. (1953) in liberal arts from the University of Richmond and an M.S. (1955) in agricultural economics from Virginia Polytechnic Institute. A member of the American Farm Economics Association, he has sought to develop more effective ways of reporting movement of products into consumption. Like Mr. Hind, he has been an agricultural economist at the USDA since 1958 and co-author of several of its publications.**

FIGURE 1
USE-THEME DISPLAY
"MAKE A WONDERFUL WALDORF SALAD TONIGHT"

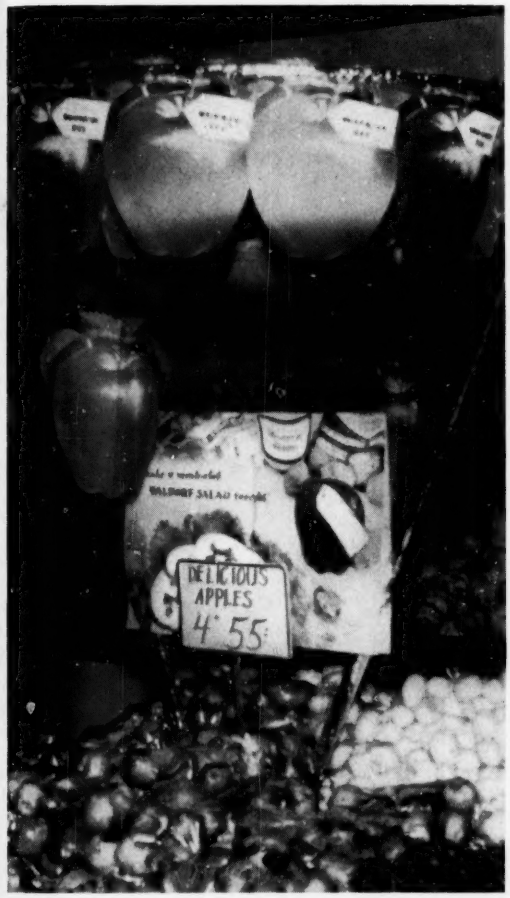



FIGURE 2
HEALTH-THEME DISPLAY
"KEEP 'EM HEALTHY"

of market simulation was described and computations made on simulated data. This paper reports the use of this design in a refined and extended form with real market data. It also discusses an entirely new technique for evaluating the effects of price, display allocation and customer traffic on the sales of advertised products.

The study was done in cooperation with the Washington State Apple Commission. During past years the Commission had developed a promotional program employing two advertising themes. One stressed the various uses of apples (fruit combination salads, baked apples, other dishes), while the other emphasized the healthful qualities of apples (builds strong bodies, dental benefits, etc.). Supermarket displays using these themes are shown in Figures 1 and 2 above.

The objectives of the study were to determine:

1. the over-all sales effectiveness of the promotional program relative to no promotion;

2. the relative sales effectiveness of the two promotional themes;

3. the short-time residual or carry-over effects, if any, of each theme;

4. the effects of the promotional activities for Washington apples on sales of apples from competing areas, and other selected fruits; and

5. the influence on sales of apples and other fruits of various merchandising and advertising practices (e.g., price cuts, display space and newspaper advertising) employed by stores to promote apples and other selected fruits.

TABLE 1

EXTRA-PERIOD LATIN SQUARE CHANGE-OVER EXPERIMENTAL DESIGN USED IN APPLE ADVERTISING STUDY OF 72 SUPERMARKETS IN SIX MIDWESTERN CITIES

| *Four-week time periods* | *Sequence (cities)* | | | | | |
|---|---|---|---|---|---|---|
| | *Square I* | | | *Square II* | | |
| | City 1 | City 2 | City 3 | City 4 | City 5 | City 6 |
| 1 | A | B | C | A | B | C |
| 2 | B | C | A | C | A | B |
| 3 | C | A | B | B | C | A |
| 4 | C | A | B | B | C | A |

Treatment A = General health theme; Treatment B = Apple use theme; Treatment C = No advertising and promotion (control group). For a discussion of the concepts underlying analysis of variance in general, and of Latin square designs in particular, please see page 40.

## METHOD

### *Experimental Design*

Items tested, called treatments, included: an apple use advertising and promotional theme, a general health advertising and promotional theme and a control with no advertising or promotion. The control treatment was used to provide a basis for comparison. The experimental design to evaluate treatment effects was an extra-period double change-over design (see Table 1). Designs of this type are sometimes referred to as switch-back, switch-over or cross-over designs.

The basic design (three time periods) consists of two replications of orthogonal Latin squares in which the sequence of treatments is reversed in the two squares. This design makes it possible to obtain estimates of direct and subsequent one-period carry-over effects of each treatment as well as the combined (direct plus carry-over) effects of each treatment with sustained use. The fourth time period, in which the treatments in the previous period are repeated, increases the precision of estimation of carry-over effects, which in turn results in greater accuracy for the estimates of the combined effects.

Treatments were assigned to successive four-week time periods (rows) and cities (columns). Since each treatment occurred once in each row and column within a basic square, systematic errors resulting from constant variations among time periods and cities were equalized.

The mathematical model which forms the basis for the analysis is:

$$Y_{ijkt} = \overline{Y} + S_i + \underline{C}_{ij} + P_{ik} + D_t + R_{t(k-1)} + e_{ijk}$$

where:

$Y_{ijkt}$ = observed sales for the $j^{th}$ city in the $i^{th}$ square during the $k^{th}$ period;

$\overline{Y}$ = over-all average sales of apples for all treatments and time periods;

$S_i$ = effect of the square;

$\underline{C}_{ij}$ = effect of the city;

$P_{ik}$ = effect of the period in that square;

$D_t$ = direct effect of treatment;

$R_{t(k-1)}$ = the carry-over (residual) effect of the immediately preceding treatment, and;

$e_{ijk}$ = experimental error.

This is the analysis of variance model, and its underlying assumptions must be reasonably met for it to be properly employed in experimental research. These assumptions are reviewed elsewhere (Eisenhart, 1947). It is sufficient to note two basic assumptions: that constants in the model can be estimated without entanglement with each other, that is, they are additive; and that the experimental errors are independently and normally distributed.

### *Sample*

Six midwestern cities, ranging in population from 100,000 to 150,000 and relatively free of sustained and intensive promotional campaigns for apples, were selected for the test. These cities were roughly comparable in major economic characteristics and supply considerations and overlapping of

their local newspaper and television facilities was negligible.

Twelve self-service food stores were selected in each city to represent establishments of different sizes, different types of management and ownership (chains, voluntary chains and nonaffiliated independents), and different geographical areas of the city. Trade sources estimated that the panel stores in each city accounted for approximately 50 to 80 per cent of retail food sales.

Intensity of promotion for the two promotional test themes (apple use and general health) featuring Washington State apples was as nearly equal as possible. Promotion included sponsored television programs Wednesday and Friday of each week and special tie-in advertising by retailers in media they normally used. During nonpromotional periods, retailers co-operated by following their normal merchandising and promotional practices for apples in the absence of a promotional campaign by a commodity group. To insure reliable measurement of advertising themes, sample stores were asked to maintain: comparable apple displays for both test promotional themes; approximately equal store-sponsored promotion of apples by each test theme (special displays and features in newspaper ads); and comparable competition from selected fruits for each test theme in price, display area and feature advertising.

Weekly tonnage sales data for Washington State apples, apples from other areas, oranges, grapefruit and bananas were collected from each store by the standard audit method:

> (Beginning inventory + weekly receipts) − (ending inventory + transfers + withdrawals + spoilage) = sales.

Additional data collected for each store on a weekly basis included total dollar sales in the produce department, and the amount of newspaper advertising by retailers for apples and selected fruits. Supplemental information on merchandising practices employed by the stores, such as prices and amount of display area for apples and other fruits, and the use of point-of-sale materials and special displays was collected by observation on Monday and Friday each week.

### *Analysis of Variance*

Analysis of variance was used to separate the variations in sales of each fruit studied (Washington State apples, all apples, oranges, grapefruit and bananas) that were attributable to cities, time periods, direct effects, carry-over effects and experimental error. For illustration, the subdivision of the sum of squares for these attributes will be shown only for sales of all apples; the subdivision of sums of squares is similar for each fruit.

In the analysis of variance, only the partitioning of the sum of squares for direct and carry-over effects of treatments require special consideration. The conventional computations for the sums of squares for the other factors stratified in the experimental design and experimental error are given in textbooks on experimental design (Federer, 1955; Lucas, 1957).

Following the notation of Table 1, the sums of squares for the direct effects of treatments is given as:

$$\frac{1}{mn(n+1)(n+2)}\left\{\left[(n+1)(\Sigma A) - (\underline{C}_2 + \underline{C}_6) - G.T.\right]^2 + \left[(n+1)(\Sigma B) - (\underline{C}_3 + \underline{C}_4) - G.T.\right]^2 + \left[(n+1)(\Sigma C) - (\underline{C}_1 + \underline{C}_5) - G.T.\right]^2\right\}$$

with m being the number of squares (two), n the common number of rows, columns and treatments (three) in a square of the basic three time-period design, and G.T., the grand total of all observations.

Obviously, since no treatments involving promotion preceded those applied in the first time period, there are no carry-over effects of treatments in the first period. Thus, in computing the sum of squared deviations associated with carry-over effects, the numerical calculations should not include the data generated in the first period. The sum of squares for carry-over effects is computed as follows:

$$\frac{1}{mn}\left[(B_{12} + A_{24} + B_{33} + C_{42} + C_{53} + A_{64})^2 + (C_{13} + C_{22} + B_{34} + B_{44} + A_{52} + A_{63})^2 + (C_{14} + A_{23} + A_{32} + B_{43} + C_{54} + B_{62})^2\right] - \frac{1}{mn^2}(G.T. - \text{total period 1})^2$$

where the first subscript for each letter indicates the city and the second subscript indicates the time period.

The treatments designed in the first set of parentheses represent all treatments immediately following treatment A in cities one through six; similarly, treatments in the second parentheses represent those following treatment B, etc.

The complete partitioning of degrees of freedom and sums of squares for total apple sales (sales of apples from all areas combined) are given in Table

2. Carry-over effects in the analysis of total apple sales did not approach statistical significance. Thus, the degrees of freedom and sums of squares associated with the carry-over effect were subsequently pooled in the error term as shown in a later table.

TABLE 2

ANALYSIS OF VARIANCE, TOTAL APPLE SALES IN POUNDS

| *Source of variation* | *Degrees of freedom* | *Sums of squares* | *Mean square* | *F* |
|---|---|---|---|---|
| Squares | 1 | 2,417,070,246 | 2,417,070,246 | 55.83* |
| Periods within squares | 6 | 2,437,012,149 | 406,168,692 | 9.38* |
| Between periods | (3) | (2,422,596,446) | (807,532,149) | 18.65* |
| Periods x squares | (3) | (14,415,703) | (4,805,234) | 0.11 |
| Cities within squares | 4 | 1,613,017,480 | 403,254,371 | 9.32* |
| Direct effect treatments | 2 | 492,948,044 | 246,474,022 | 5.69† |
| Carry-over effect treatments | 2 | 7,062,739 | 3,531,370 | .08 |
| Error | 8 | 346,325,670 | 43,290,709 | — |
| Total | 23 | 7,313,436,328 | — | — |

Under certain experimental conditions, the statistician might argue to pool the sums of squares and degrees of freedom for period x squares interaction (when insignificant) in the error term. However, when to pool and when not to pool is controversial and may tend to inflate the significance of treatment effects as in this case. Thus, for conservative estimates of treatment effects, this was not done.

* Significant at the .01 probability level.

† Significant at the .05 probability level.

When significant carry-over effects are present, the following computational procedures are employed for estimating the direct and carry-over effect of each treatment.

Estimates of the direct effect per treatment, $\overline{D_a}$ $\overline{D_b}$ $\overline{D_c}$ are:

$$\overline{D_a} = \overline{Y} + \frac{1}{mn(n+2)}\left[n+1\,(\Sigma\,A) - (\underline{C_2} + \underline{C_3}) - G.T.\right]$$

$$\overline{D_b} = \overline{Y} + \frac{1}{mn(n+2)}\left[n+1\,(\Sigma\,B) - (\underline{C_3} + \underline{C_1}) - G.T.\right]$$

$$\overline{D_c} = \overline{Y} + \frac{1}{mn(n+2)}\left[n+1\,(\Sigma\,C) - (\underline{C_1} + \underline{C_2}) - G.T.\right]$$

Estimates of carry-over effects per treatment, $\overline{R_a}$, $\overline{R_b}$, $\overline{R_c}$, are:

$$\overline{R_a} = \frac{1}{mn^2}\left[n(B_{12} + A_{24} + B_{23} + C_{42} + C_{53} + A_{34}) - (G.T. - \text{Total period 1})\right]$$

$$\overline{R_b} = \frac{1}{mn^2}\left[n(C_{13} + C_{22} + B_{34} + B_{44} + A_{52} + A_{63}) - (G.T. - \text{Total period 1})\right]$$

$$\overline{R_c} = \frac{1}{mn^2}\left[n(C_{14} + A_{23} + A_{32} + B_{43} + C_{54} + B_{62}) - (G.T. - \text{Total period 1})\right]$$

with symbols defined as before.

The estimate of the combined and direct carry-over of a particular treatment is the sum of the means of the direct and carry-over effects for that treatment, $\overline{D_t} + \overline{R_t}$, computed by the preceding formulae. Since advertising investment is usually spread over a sustained period of time, the combined direct and carry-over effects of treatments would appear to provide better estimates of treatment differences than direct effects alone (i.e., when carry-over effects are significant).

For treatment contrasts, the standard error of a difference for direct, carry-over, and combined (direct plus carry-over) effects is given respectively as the square root of:

$$\frac{2(n+1)}{mn\,(n+2)}.S^2,\ \frac{2}{mn}.S^2,\text{ and }\frac{2(2n+3)}{mn\,(n+2)}.S^2$$

where $S^2$ is the error mean square obtained in the analysis and m and n are as defined previously.

In the illustrative study, only the direct effects of treatments are considered, since carry-over effects were negligible. Estimates of direct effects of treatments, and the accompanying sample statistics, are given in Table 3 at this stage of the analyses (i.e., before covariance analysis).

TABLE 3

TREATMENT MEANS AND SAMPLE STATISTICS BEFORE AND AFTER COVARIANCE ANALYSIS[1]

| *Item* | *Before covariance analysis* — *Direct effect* ($\overline{D_t}$) | *After covariance analysis* — *Direct effect* ($\overline{D_t}$) |
|---|---|---|
| Treatment means: (Direct effects) | | |
| Apple use theme | 75,461 | 77,267 |
| General health theme | 71,575 | 69,722 |
| No advertising or promotion | 64,176 | 64,222 |
| Sample statistics: | | |
| Standard deviation[2] | 5,945 | 5,333 |
| Coefficient of variation (%)[2] | 8.4 | 7.6 |
| Standard error of a difference[3] | 3,070 | 2,754 |

[1] Treatment means are expressed as average sales per city per four-week time period.

[2] Based on pooled degrees of freedom and sums of squares for carry-over effects in the experimental error.

[3] Between two treatment means.

### *Analysis of Covariance*

The sales data for Washington State apples and all apples were also adjusted for nonconstant sources of variation in sales that were not taken into account by the experimental design. These included

variation in weather conditions among cities within time periods. The store's total dollar sales in the produce department were used in covariance analysis as an index of number of customers and purchasing power of customers in making adjustments. Produce sales did not completely satisfy the requirements of a concomitant observation, because produce sales and apple sales were not independent of each other since produce sales are affected to some extent by apple sales. However, apple sales contribute a relatively small percentage of the store's total produce sales. This was the best index for which data were available to reflect customer traffic and purchasing power.

The covariance correction for the regression of sales of all apples on total produce sales increased the precision of the findings because results were then based upon a constant number of customers and customer purchasing power in each city during each treatment. A brief description of the mathematical computations follows.

The analysis of sums of squares of produce sales, and the sums of products of produce sales and the sales of apples, is analogous to the analysis of variance previously described. The only difference is that in computing the cross products corresponding values of produce and apple sales are multiplied instead of squared in each stage of computation.

If $Y_{ijkt}$ is defined as before and $X_{ijkt}$ is the corresponding concomitant observation for produce sales and b is the constant multiplier for the deviations of the concomitant variable from its over-all mean, the previous analysis of variance model takes the form:

$$Y_{ijkt} = \overline{Y} + S_i + \underline{C}_{ij} + P_{ik} + D_t + D_{t(k-1)} + b\,(X_{ijkt} - \overline{X}) + e_{ijk},$$

with the symbols as defined before and $\overline{X}$ the over-all mean of produce sales.

The effects of the regression of produce sales (X) are removed from the sums of squares for error and direct effects treatment + error of apple sales (Y) by using the respective sums of squares for X and Y and corresponding cross products (XY) shown in Table 4 in the formula:

$$S_Y^2 - (S_{XY})^2 / S_X^2$$

A degree of freedom is associated with regression and subtracted from degrees of freedom for error and treatment + error. The sum of squares of direct effect treatments, adjusted for produce sales, are then obtained by subtracting the corrected sum of squares for error from the corrected sum of squares for treatment + error. Degrees of freedom for treatments are similarly obtained. The complete computations can be followed in Table 4. A more detailed description of the standard computations involved are given in *Biometrics* (Cochran, 1957).

Adjusted treatment means are computed by: $\overline{D}_t - b\,(\overline{X}_t - \overline{X}_{ijk})$ where $\overline{D}_t$ is the mean apple sales of the $t^{th}$ treatment, $\overline{X}_t$ is the corresponding mean for produce sales and $\overline{X}_{ijk}$ is the over-all mean of produce sales. The adjustment factor or regression coefficient b is given by the formula, $b = \frac{S_{XY}}{S_X^2}$ where $S_{XY}$ and $S_X^2$ are the error sums of cross products (of X and Y) and the error sum of squares (for X) respectively, as shown in Table 4.

Adjusted treatment means, and other sample statistics computed before and after the covariance analysis are shown in Table 3. The increase in accuracy due to covariance analysis is demonstrated by the ten per cent reduction in the size of the

TABLE 4

ANALYSIS OF COVARIANCE, TOTAL APPLE SALES IN POUNDS (Y), PRODUCE SALES IN DOLLARS (X)

| *Source of variation* | *Degrees of freedom* | *Sum of squares and products* | | | *Error of estimate*[1] | | |
|---|---|---|---|---|---|---|---|
| | | $S_{X^2}$ | $S_{XY}$ | $S_{Y^2}$ | *Sum of squares* | *Degrees of freedom* | *Mean square* |
| Direct effects of treatments | 2 | 36,021,777 | 42,438,284 | 492,948,044 | — | — | — |
| Error[2] | 10 | 69,926,512 | —82,535,043 | 353,388,409 | 255,971,376 | 9 | 28,441,264 |
| Treatments + error | 12 | 105,948,289 | —40,096,759 | 846,336,453 | 831,161,596 | 11 | — |
| Adjusted for produce sales | — | — | — | — | 575,190,220 | 2 | 287,595,110[3] |

[1] Sum of Squares for error: $S_{y^2} - \frac{(S_{XY})^2}{S_{X^2}} = 353{,}388{,}409 - \frac{(-82{,}535{,}043)^2}{69{,}926{,}512} = 255{,}971{,}376$

Sum of Squares for treatments + error: $S_{y^2} - \frac{(S_{XY})^2}{S_{X^2}} = 846{,}336{,}453 - \frac{(-40{,}096{,}759)^2}{105{,}948{,}289} = 831{,}161{,}596$

[2] Carry-over effects are insignificant and are pooled in error terms.

[3] This treatment mean square with two degrees of freedom tested against error mean square with nine degrees of freedom is significant at the .005 per cent probability level.

standard error of a difference. The use of the concomitant variable in the covariance analysis had almost the same effect as an additional Latin square in reducing the size of the difference between two treatment means which would be statistically significant. That is, at least one more replication of the four-period Latin square design would be required to attain the same precision in detecting significant differences if covariance were not used.

The F test was used to determine if differences in sales among the three treatment means were statistically significant. The least significant difference (LSD) test was used to determine significance between any two treatment means. These tests will not be discussed in detail, since they are commonly used methods (Cochran, 1957). It should be noted, however, that a prerequisite to using the LSD test is the finding of significant differences among treatments by the F test.

### *Multiple Covariance Analysis*

An extension of the covariance technique discussed previously was used to determine the nature and extent of the influence of retail merchandising practices and pricing policies on sales of apples (Washington State and other areas), oranges, grapefruit and bananas. Merchandising factors evaluated were price, display space and newspaper advertisement space. Produce sales, as before, were used to reflect number of customers and relative purchasing power of customers. These merchandising factors are referred to hereafter as quantitative factors. It was not possible to measure the direct effects on sales of such nonquantitative factors as variety, size and quality of fruits, size of pricing unit, type of display (prepackaged, bulk or combination) and packaging material.

Data for quantitative factors were tabulated for each city by weeks and plotted on scatter diagrams against sales of each fruit. Thus, a general indication was obtained of the factors related to volume of sales of each fruit. Factors which had no apparent relation to sales were eliminated.

Arithmetical computations in multiple covariance analysis are most laborious and complex. Ostle (1954) reviews them briefly. It would be quite boring and rather lengthy to discuss and illustrate the many iterative computations involved. These analyses are better explained by symbolizing a basic regression model and then following with some discussion of the adaptations and extensions of this basic model as used in the study.

The basic regression model is:

$$Y = \mu + b_1x_1 + b_2x_2 + \ldots + b_nx_n + e$$

where Y is the sales observation in a city during any week for a particular fruit, $\mu$ the over-all mean and $b_1$, $b_2$ ... $b_n$ regression coefficients, estimates of the effects of the $x_1$, $x_2$ ... $x_n$ selected quantitative factors. The residual e is made up of city-time period and treatment effects, and random or experimental error. The regression model assumes that such quantitative factors as price, display area and newspaper advertisement space for the particular fruit remain constant between cities and four-week time periods. That is, none of the variation in sales is due to the influence of other variables associated with city and time period differences such as income levels, population characteristics and seasonal trends. This is highly unrealistic, since the influence of such variables vary considerably between cities and over time.

However, by combining the concepts of regression and analysis of variance, a more discriminating analysis can be made in which place and season effects are removed. This technique, covariance analysis, permits the measurement of the net effects of these specified quantitative factors on sales. The analysis of variance model is:

$$Y_{ij} = \mu + P_i + C_j + e_{ij}$$

where $Y_{ij}$ and $\mu$ are defined as before, $P_i$ and $C_j$, the city and time period respectively. In this model the residual $e_{ij}$ consists of the effects of the quantitative factors $b_1x_1 + b_2x_2 + \ldots b_nx_n$ and the random noncompensating errors of measurement. Combining the regression and analysis of variance models we have the covariance model:

$$Y_{ij} = \mu + P_i + C_j + b_1x_1 + b_2x_2 + \ldots b_nx_n + z_{ij}$$

where $z_{ij}$ represents the random and noncompensating errors and the effects of the unidentified factors. This model, unlike the previously stated regression model, defines and accounts for the effects of city and seasonal differences. Thus, the estimates of the effects $b_1$, $b_2$ ... $b_n$ of the quantitative factors $x_1$, $x_2$, ... $x_n$ on sales are free of the place and season effects.

Based on the covariance model, a multiple analysis of covariance was first used to adjust the sales variation for each fruit for the variations associated with cities and time periods. A multiple regression analysis was then made of the adjusted data (i.e., the residual sums of squares and cross products) to identify and quantify the net effects of the merchandising factors significantly affecting sales. The

TABLE 5

MULTIPLE COVARIANCE ANALYSIS

| *Source of variation* | *Degrees of freedom* | *Sums of squares* | *Mean squares* | *F*[1] |
|---|---|---|---|---|
| Total | 863 | 412,284,650 | — | — |
| Nonquantitative factors: | | | | |
| Cities and time periods | 17 | 129,383,870 | 7,610,816 | 44.8 |
| Residual 1 | 846 | 282,900,780 | 334,398 | — |
| Quantitative factors:[2] | | | | |
| Display space for Washington State apples | 1 | 48,705,457 | 48,705,457 | 286.6 |
| Produce sales[3] | 1 | 41,561,786 | 41,561,786 | 244.5 |
| Display space for other apples | 1 | 4,861,877 | 4,861,877 | 28.6 |
| Price of Washington State apples | 1 | 3,630,474 | 3,630,474 | 21.4 |
| Display space for grapefruit | 1 | 2,558,838 | 2,558,838 | 15.1 |
| Newspaper advertisement space for Washington State apples | 1 | 2,113,867 | 2,113,867 | 12.4 |
| Joint effects of above factors (interactions) | 4 | 36,697,381 | — | — |
| Residual 2 | 840 | 142,771,100 | 169,996 | — |

[1] All significant at the 0.01 per cent probability level.
[2] Quantitative factors adjusted for effects of each other and nonquantitative factors of cities and time periods.
[3] Used as an index to reflect changes in the number of customers patronizing the store and changes in the relative purchasing power of customers.
[4] Degrees of freedom entangled with residual degrees of freedom.

multiple regression analysis was repeated until only those factors affecting sales remained which attained statistical significance at the 0.05 probability level. The practical utility of a factor based upon the magnitude of its regression coefficient (b value) and coefficient of determination ($R^2$) was also a criterion for retaining it in subsequent analysis. The complete analysis for Washington State apple sales is shown in Table 5.

## RESULTS

There were substantial differences in sales of both Washington State and all apples between periods with promotional themes (apple use and health) and periods of no promotion (see Table 6).

When sales of apples were combined, the apple use theme was significantly more effective in promoting sales than the health theme. However, the nine per cent sales difference between the two themes for Washington State apples was not large enough to be statistically significant.

The themes used in the four-week test period significantly affected neither Washington State *nor* total apple sales during the next four-week period without advertising.

TABLE 6

SALES OF APPLES FROM WASHINGTON STATE AND OTHER AREAS DURING NO APPLE PROMOTION AND TWO APPLE PROMOTIONAL THEMES[1]

| *Source of apples* | *Average sales per store per four-week period* | | | *Difference in sales between no promotion and—*[2] | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | *With no promotion* | *With apple-use theme* | *With health theme* | *Apple-use theme* | | *Health theme* | | *Apple-use theme and health theme* | |
| | *Lb.* | *Lb.* | *Lb.* | *Lb.* | *%* | *Lb.* | *%* | *Lb.* | *%* |
| Washington State | 3,124 | 4,117 | 3,784 | 993** | 31.8 | 660* | 21.2 | 333 | 8.8 |
| Other areas (Midwestern and Eastern States and Canada) | 2,227 | 2,322 | 2,026 | 95 | 4.3 | —201 | —9.1 | 295 | 14.6 |
| Total or average | 5,351 | 6,439 | 5,810 | 1,088*** | 20.1 | 459* | 8.6 | 629** | 10.8 |

[1] All sales data were adjusted for variations among treatments (themes and no promotions) which might be attributed to differences in number of customers and purchasing power per customer. While there were price reductions associated with weekly specials featured by retailers, average prices (weighted by sales) did not differ significantly among treatments.
[2] Differences required for statistical significance for apples from Washington State and from other areas are ± 735 pounds at the five per cent probability level**; ± 596 pounds at the 10 per cent probability level*. Differences required for statistical significance for apples from all areas combined are ± 748 pounds at the one per cent probability level***; ± 520 pounds at the five per cent probability level**; and ± 422 pounds at the 10 per cent probability level*.
The economic implication of these probability levels is that the chances or odds of obtaining sales differences of these magnitudes are less than 1, 5 or 10 in 100 from random sales flucuations and non-constant experimental errors.

TABLE 7

SALES OF SELECTED FRUITS DURING NO APPLE PROMOTION AND TWO APPLE PROMOTIONS[1]

| Fruit | Average sales per store per four-week period | | | Difference in sales between no promotion and—[2] | |
|---|---|---|---|---|---|
| | With no promotion | With apple use theme | With health theme | Apple use theme | Health theme |
| | Lb. | Lb. | Lb. | % | % |
| Oranges | 5,516 | 5,680 | 5,784 | 3.0 | 4.9 |
| Grapefruit | 5,272 | 5,156 | 5,996 | —2.2 | 13.7 |
| Bananas | 5,944 | 5,836 | 5,752 | —1.8 | —3.2 |
| Total | 16,732 | 16,672 | 17,532 | —0.4 | 4.0 |

[1] Ali sales data were adjusted for variations among advertising treatments which might be attributed to differences in number of customers and purchasing power per customer. Sales data were further adjusted for the effects of prices, display space, and other significant merchandising and promotional practices employed by stores. The regression coefficients obtained in the multiple regression analysis were used in adjusting means. Computations are similar to those made in obtaining the adjusted means in the covariance analysis.

[2] Differences required for statistical significance at the five per cent probability level are ± 10.8 per cent for oranges; ± 17.5 per cent for grapefruit; and ± 10.0 per cent for bananas. Differences required for statistical significance at the 10 per cent probability level are ± 8.7 per cent for oranges; ± 14.2 per cent for grapefruit; and ± 8.1 per cent for bananas.

Advertising Washington State apples exerted only a minor influence on sales of oranges, grapefruit and bananas. The effect of the advertising seemed to vary among the fruits depending on the theme employed. The sales differences were too small to determine whether the promotional themes for apples significantly improved the sales of these fruits or not. The differences that were found corresponded to findings of previous research studies, namely, that advertising and merchandising practices which increase sales of apples also benefit sales of oranges, as suggested by the data in Table 7. Also, the decrease in sales of bananas when either apple theme was advertised compared to no promotion was similar to findings of previous studies which have indicated that apples and bananas are competitive products (Henderson, 1952, 1955, 1955).

Changes in sales of apples, oranges, grapefruit, and bananas were significantly related to changes in some but not all of the practices employed by stores in merchandising and promoting these fruits, such as amount of display area, newspaper advertising, and prices (see Table 8).

Sales of each fruit were generally affected by the merchandising and promotional practices used with it. The major exception was the amount of display space devoted to each fruit, which affected the fruit displayed and also had some influence on other fruits. Variation in the amount of display space used for grapefruit affected all fruit except bananas. Grapefruit sales varied directly with the amount of space in grapefruit displays, while sales of apples and oranges varied inversely with the amount of space for grapefruit displays. Banana sales increased and orange sales decreased when the size of banana displays was increased. Similarly, a decrease in banana sales and an increase in orange sales were associated with a decrease in the amount of space devoted to banana displays. These findings indicate that grapefruit competes with apples and oranges for display space and sales, while bananas compete with oranges for display space and sales.

Price and display space devoted to each fruit exerted the most influence on sales. Sales for each fruit varied directly with the amount of display space and inversely with price. The variation in sales of each fruit from week-to-week was also generally related to the week-to-week variation in amount of newspaper advertising space devoted to each fruit by retailers.

Sales of each fruit were significantly and directly related to the volume of produce sales. Produce sales reflect the combined effects of promotional and merchandising practices employed by stores on the sales of individual products, and the influence of such practices in drawing additional customers into the stores.

## OTHER APPLICATIONS

While the notation and terminology used refer to this specific experiment, the design and analyses presented would be valid in other areas of advertising research where measurable carry-over effects of promotional techniques are likely to occur. Some general applications are given in the following illustrations.

This experimental design could be used to predict the most efficient promotional alternatives from consumer or trade media advertising, personal selling, point-of-purchase effort, cooperative advertising and premium offers, and to determine the place of each in the total promotional effort for a

TABLE 8

MERCHANDISING AND PROMOTIONAL PRACTICES WHICH SIGNIFICANTLY AFFECTED SALES OF WASHINGTON STATE APPLES, APPLES FROM OTHER AREAS, AND OTHER SELECTED FRUITS

| Factors | Factors significantly affecting sales of—[1] | | | | |
|---|---|---|---|---|---|
| | *Washington State apples* | *Other apples*[2] | *Oranges* | *Grapefruit* | *Bananas* |
| Produce sales (dollars) | + 0.3 | + 0.3 | + 0.7 | + 0.8 | + 0.7 |
| Price: (cents per pound) | | | | | |
| Washington State apples | —25.9 | — | — | — | — |
| Other apples | — | — 8.8 | — | — | — |
| Oranges | — | — | —93.9 | — | — |
| Grapefruit | — | — | — | —122.1 | — |
| Bananas | — | — | — | — | —86.6 |
| Display space: (square feet) | | | | | |
| Washington State apples | +21.8 | — 8.4 | — | — | — |
| Other apples | — 6.8 | +25.8 | — | — | — |
| Oranges | — | — | +16.2 | — | — |
| Grapefruit | — 4.8 | — 6.4 | — 7.3 | + 10.7 | — |
| Bananas | — | — | —14.9 | — | +15.6 |
| Newspaper advertising space: (square inches) | | | | | |
| Washington State apples | + 5.0 | — | — | — | — |
| Oranges | — | — | + 6.5 | + 8.5 | — |
| Grapefruit | — | — | — | — | — |
| Bananas | — | — | — | — | +19.6 |

[1] A plus sign indicates that a positive change (increase) in the value of a factor is accompanied by an increase in sales of a fruit and a negative change by decrease in sales. For example, on the average, an increase of one square foot in the display space for Washington State apples was accompanied by an increase of 21.8 pounds in Washington apple sales. A negative sign signifies that a positive change in the factor results in decreases in Washington apple sales and a negative change (decrease) in the factor results in an increase in Washington apple sales. Thus, on the average an increase of one cent a pound in the price of Washington apples resulted in a decrease of 25.9 pounds in sales of Washington apples.

[2] Apples from areas other than Washington State.

product. At the same time, holding promotional expenditures relatively constant for each technique tested, promotional outlay could be related to sales returns. Using a more complicated arrangement of this design, the optimum levels of promotional outlay under different conditions could also be found.

With these alternative uses of the design the multiple covariance technique could be used to obtain estimates of sales responses (for a product sold at retail) to changes in prices, display space and other merchandising practices employed by retail stores. This information could be used in deciding what combination of price and merchandising practices will tend to maximize returns from advertising and promotional activities. The covariance technique could also determine relationships between sales of a product and such factors as price of the product, price of related or competing products, and consumers' incomes, in order to forecast sales in a market from measurements on a sample. The estimated parameters for such quantitative variables could be applied to other market areas with similar socio-economic characteristics, since these estimates would be corrected for place and season effects.

REFERENCES


COCHRAN, W. G. et al. A Double Change-over Design for Dairy Cattle Feeding Experiments. *Journal of Dairy Science*, Vol. 24, No. 11, November 1941, pp. 937-951.

COCHRAN, W. G. Analysis of Covariance: Its Nature and Uses. *Biometrics*, Vol. 13, No. 3, September 1957, pp. 268-269.

COCHRAN, W. G. *Experimental Designs*. Second Edition. New York: John Wiley & Sons, Inc., 1957, p. 76.

EISENHART, CHURCHILL. The Assumptions Underlying the Analysis of Variance. *Biometrics*, Vol. 3, No. 1, March 1947, pp. 1-21.

FEDERER, WALTER T. *Experimental Design*. New York: The Macmillan Company, 1955, pp. 148-149 and 449-452.

HENDERSON, PETER L. Application of the Double Change-over Design to Measure Carryover Effects of Treatments in Controlled Experiments. *Methods of Research in Marketing. Paper No. 3*. Ithaca: Cornell University, 1952.

HENDERSON, PETER L. *Marketing Services Affecting Apple Sales*. Unpublished Ph.D. Dissertation. Ithaca: Cornell University Library, 1952.

HENDERSON, PETER L. Merchandising Apples. *Virginia Fruit*. Virginia Horticultural Society, Vol. 42, May, 1955, pp. 34-36.

HENDERSON, PETER L. The Effect of Advertising and Point of Sale Aids on Apples. *Virginia Fruit*, Vol. 42, June 1955, pp. 30-34.

JESSEN, RAYMOND J. A Switch-over Experimental Design to Measure Advertising Effect. *Journal of Advertising Research*, Vol. 1, No. 3, March 1961, pp. 15-22.

LUCAS, H. L. Extra-period Latin Square Change-over Design. *Journal of Dairy Science*, Vol. 40, No. 3, March 1957, pp. 225-239.

OSTLE, BERNARD. *Statistics in Research*. Ames: The Iowa State College Press, 1954, pp. 406-408.

WOOD RESEARCH CORPORATION. The Experimental Method in Marketing Research. Philadelphia: *Wood Chips*. February 1959.

# A Sensitive Measure of Ad Exposure

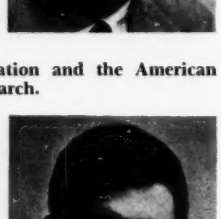
GORDON M. KESWICK and LAWRENCE G. COREY

Needham, Louis and Brorby, Inc.

**The authors have developed a simple apparatus for measuring ad recognition, and found it discriminated among advertisements to which the reader was exposed much, little or not at all.**

THIS PAPER DESCRIBES a new technique for measuring reader recognition of print advertisements. The system, called the Controlled Ad Awareness Technique (CAAT), shows ads to the respondent through one or more screens. This procedure aids recall, restricts visibility of the advertisement's graphic and copy content, and permits systematic presentation of increasing amounts of stimulus information. The CAAT neither invites the false recognition claims inherent in the Starch method, nor makes the unreasonable demands on memory of Gallup & Robinson's aided recall measure.

**GORDON M. KESWICK is director of creative research at Needham, Louis and Brorby, Inc. in Chicago. He received a Ph.D. in communication research from the University of Illinois in 1958. Before joining Needham, Louis and Brorby in 1957 he worked in data processing and presentation at the Control Systems Laboratory, University of Illinois. He has lectured at the Institute of Design, the Illinois Institute of Technology and to special advertising groups. He is a member of the American Psychological Association and the American Association for Public Opinion Research.**



**LAWRENCE G. COREY is a research associate at Needham, Louis and Brorby. Previously he served as associate director of aging and retirement studies at the Family Study Center of the University of Chicago. Holder of a B.A. from Roosevelt University (1956) and an M.A. from the University of Chicago (1958), he has published research on personality measurement in the *Journal of Gerontology* and the *Journal of Applied Psychology*.**

The Starch noted score, though reliable (Lucas, 1960), is seriously inflated by influences extraneous to actual ad exposure. Appel and Blum (1961) found that a respondent was more likely to claim having noted an unpublished ad if he was interested in the product advertised, about to purchase it, familiar with the editorial content of the magazine, and a reader of many magazines.

At the other extreme, Gallup & Robinson's aided recall score makes undue demands on the respondent's memory. As Lucas put it, "The [Gallup & Robinson] qualifying procedure and the interview approach do not provide a maximum background of legitimate cues for remembering advertisements" (1960, p. 17).

FIGURE 1. HOW A PHOTOGRAPH APPEARS:

THROUGH ONE SCREEN

WITH NO SCREEN

Compared with these methods, the CAAT seems to have at least three advantages:

1. It minimizes—perhaps eliminates—the effects on recognition of factors other than actual exposure to advertisements, and progressively increases partial stimulus information as an aid to recall.
2. It measures *degrees* of recognition in contrast to the "all or nothing" recognition dichotomy characteristic of the Starch and Gallup & Robinson measures.
3. It can evaluate with equal facility reader recognition of advertisements which have appeared only once, and accumulated recognition resulting from any number of exposures to the same advertisement.

In the Controlled Ad Awareness Technique, the interviewer exposes advertisements placed between the innermost screen and a heavy cardboard backing, in successive discrete steps. There are many ways in which control can be imposed. In our first study, we used three Zip-A-Tone 080-R opaque plastic screens, essentially identical, and each embossed with a mottled white, random pattern. All three screens were attached, dull side up, to a cardboard backing, such that they could be lifted one at a time, at the experimenter's discretion.

A test ad was inserted between the screens and the backing. Respondents were given four successive opportunities to identify it: through three screens, through two screens, through one, and unscreened. When all three screens covered an ad, its visual cues were minimal; as the first and second screens were lifted, more and more visual information appeared.

Figure 1 shows a conventional photograph unscreened and rephotographed through one screen. While the latter appears somewhat different from the picture as actually seen through one screen, it does convey the effect. Ads in which respondents were strongly interested and which, consequently, were more closely attended, might well be recognized through greater densities of screening than will ads which were merely glanced at. A totally unfamiliar ad probably would be impossible to identify through the interposed screens.

The method produces a recognition score which corresponds to the number of screens through

which the ad was identified. Thus the instrument generated an ordinal recognition scale which permits a broader range of statistical treatment than do the Starch or Gallup & Robinson scores of the per cent who note or recall.

## PURPOSE

The preceding suppositions about the sensitivity and workability of the CAAT were tested in August 1961. The major objective of the research was to measure the CAAT's sensitivity under controlled conditions.

Using ads which respondents could not have seen previously, we asked whether exposure for carefully controlled periods of time, and under different states of induced interest, would produce expected differences in identification at given levels of screening. Specifically, this hypothesis was tested: the number of screens through which advertisements are correctly identified will increase as a direct function of increasing levels of pre-exposure and induced interest in the product or service advertised.

This hypothesis was tested using a sample of 50 Needham, Louis and Brorby employees, 21 men and 29 women. Besides their being available, agency personnel seemed suitable as subjects for two reasons. First, the advertisements selected for the study were totally unfamiliar to the subjects as they were for West Coast regional brands and services and had appeared only in a West Coast magazine. Second, there was no reason to suppose that our employees would be significantly different from anyone else in the task of identifying screened advertisements out of the context of their agency responsibilities.

## METHOD

The study was done in two phases: Phase I, where subjects were exposed to test and buffer advertisements systematically; and Phase II, in which subjects were asked to identify exposed and unexposed test ads inserted in the CAAT screening device.

### *Phase I: Exposure*

Three degrees of controlled exposure to eight unfamiliar print ads served as the criteria against which sensitivity of the CAAT was measured. Each level of exposure was prefaced by standard instructions covering degree of product interest and the amount of time respondents were to invest in particular ads. Levels were:

*Maximum* exposure. Thorough reading and examination of the illustrations in the advertisements as though respondents were strongly and immediately interested in the advertised product.

*Moderate* exposure. Thorough examination of the illustration and a reading of the headline but no more. Interest at this level was to be in the eye appeal of the illustration rather than in the product being advertised.

*Minimum* exposure. A glance at the total advertisement as the page was flipped.

Certain other test ads were *not* exposed to respondents at all.

Eight full-page test ads, plus eight buffer advertisements, equally divided between color and black and white, were selected from a West Coast magazine. Six test and buffer ads were combined into four different portfolios of 12 ads each, so that different pairs of test and buffer ads were absent from each of the four portfolios. Each of the 16 ads received three levels of pre-exposure: maximum, moderate, minimum—and no exposure. Also, the order in which the ads were pre-exposed to subjects was changed from group to group.

Table 1 lists the previously defined print ad exposure levels by portfolio (excluding the buffer ads which were not put to the test of screened identification).

TABLE 1

EXPOSURE LEVELS OF TEST ADS

| *Portfolio* | *Maximum* | *Moderate* | *Minimum* | *No Exposure* |
|---|---|---|---|---|
| I | Furniture*<br>United Bank | Dishwashers*<br>Tomato juice | Motorama*<br>Free prints | Wells Fargo*<br>Egg |
| II | Wells Fargo<br>Egg | Furniture<br>United Bank | Dishwasher<br>Tomato juice | Motorama<br>Free prints |
| III | Motorama<br>Free prints | Wells Fargo<br>Egg | Furniture<br>United Bank | Dishwasher<br>Tomato juice |
| IV | Dishwasher<br>Tomato juice | Motorama<br>Free prints | Wells Fargo<br>Egg | Furniture<br>United Bank |

* Black and white print ad.

### *Phase II: Recognition*

In individual interviews, all eight test ads, including the pairs absent from each of the four portfolios, were presented to subjects in the CAAT device between 24 and 72 hours after pre-exposure.

The eight ads were inserted in the CAAT device under the innermost screen, and removed one at a time following successful or unsuccessful identification. The order of the ads was systematically rotated from subject to subject. In these sessions subjects were not permitted to touch the screening device; the interviewer smoothed and lifted the screens as his questioning proceeded.

Subjects were told that they were going to be shown eight ads, some they had seen in Phase I and others they had not, to identify the general product class as soon as they could and to tell the interviewer whether they recalled having seen the ad during the previous exposure session.

## RESULTS

Rank correlations between exposure and identification levels were the same for men and women, and for subjects taking the CAAT 24 and 72 hours after pre-exposure. Consequently all subjects were analyzed as a single population, regardless of sex and time lapse differences.

The CAAT would be considered sensitive, and our working hypothesis confirmed, if a majority of the test ads generated high positive coefficients of correlation between scores in Phases I and II (i.e., maximum exposure/maximum identification, moderate exposure/moderate identification, etc.). To be considered significant, a correlation coefficient had to differ from zero at or beyond the .001 level of confidence.

Table 2 shows identifications of one ad by 50 respondents. Note that the levels of correct screened identification increased as a direct function of increasing levels of exposure and induced interest.

TABLE 2

EXPOSURE AND IDENTIFICATION OF FURNITURE AD (Rho = .82, N = 50)

| | *Identification Level* | | | | |
|---|---|---|---|---|---|
| *Exposure* | *3 Screens* | *2 Screens* | *1 Screen* | *No Screen* | *Total* |
| Maximum | 5 | 7 | 0 | 0 | 12 |
| Moderate | 4 | 4 | 3 | 0 | 11 |
| Minimum | 0 | 3 | 11 | 0 | 14 |
| None | 0 | 0 | 7 | 6 | 13 |
| Total | 9 | 14 | 21 | 6 | 50 |

Table 3 summarizes the exposure-recognition correlations for all eight ads. The median of these coefficients is .715 and all but the last were significant at the .001 level of confidence.

TABLE 3

CORRELATIONS OF EXPOSURE WITH IDENTIFICATION LEVELS (N = 50)

| *Print Ad* | *Rho* |
|---|---|
| Furniture* | .82 |
| Free prints | .80 |
| Motorama* | .77 |
| Egg | .73 |
| Tomato juice | .70 |
| Dishwasher* | .62 |
| United Bank | .60 |
| Wells Fargo* | .38 |

* Black and white ad.

The CAAT's sensitivity is also shown in the average identification scores in Table 4. These scores are graphed in Figure 2 on the next page. Note that the more that ad was exposed, the more screens through which it was correctly identified.

TABLE 4

MEAN NUMBER OF SCREENS THROUGH WHICH THE AD COULD BE IDENTIFIED (N = 50)

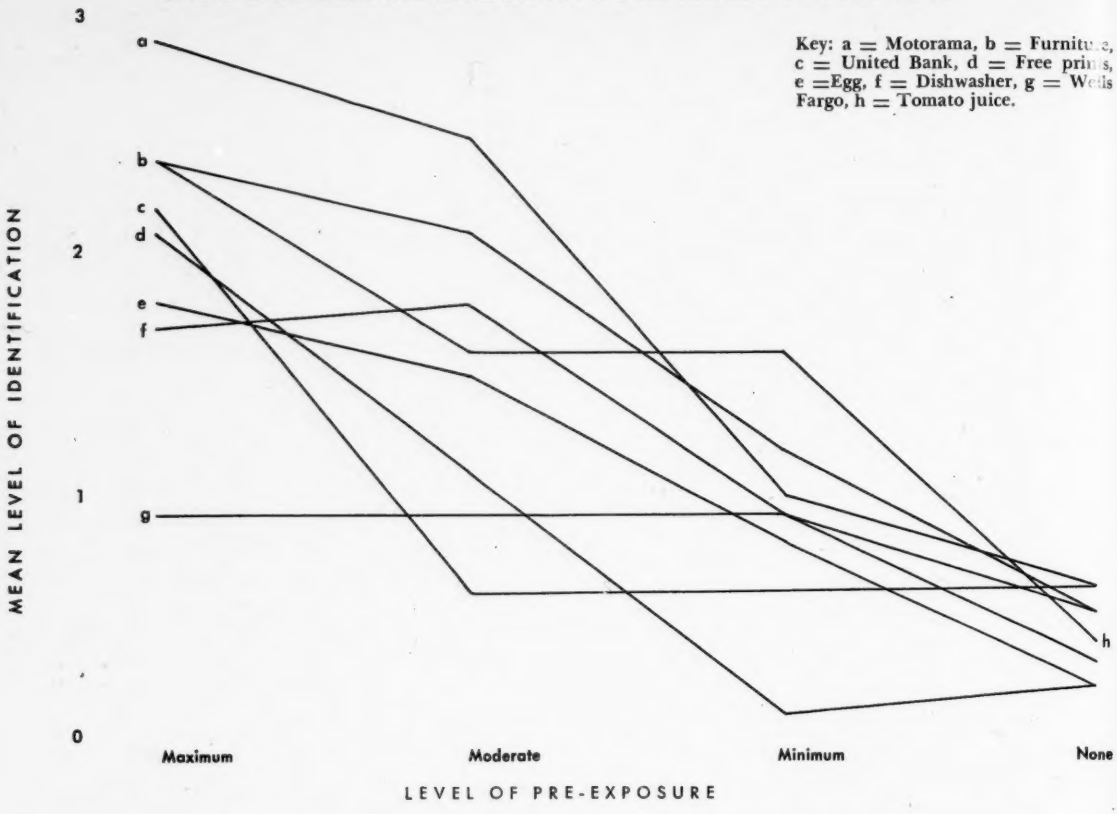
| | *Level of Exposure* | | | |
|---|---|---|---|---|
| *Print Ad* | *Maximum* | *Moderate* | *Minimum* | *None* |
| Furniture* | 2.4 | 2.1 | 1.2 | 0.5 |
| Free prints | 2.1 | 1.1 | 0.1 | 0.2 |
| Motorama* | 2.9 | 2.5 | 1.0 | 0.6 |
| Egg | 1.8 | 1.5 | 0.8 | 0.2 |
| Tomato juice | 2.4 | 1.6 | 1.6 | 0.4 |
| Dishwasher* | 1.7 | 1.8 | 0.9 | 0.5 |
| United Bank | 2.2 | 0.6 | 0.6 | 0.6 |
| Wells Fargo* | 0.9 | 0.9 | 0.9 | 0.3 |

* Black and white ad.

## DISCUSSION

As predicted, the CAAT discriminated among ads according to their prior exposure. These results, while convincing, are based on only 50 respondents. We must entertain the possibility that these respondents may have been somewhat biased by their unusual sensitivity to advertising. The most serious consequence of respondents' sensitivity to advertising would be an upward shift of their identification scores from those expected from the general population. This would not suggest rejection of our hypothesis, however, since there is no reason to believe that such upward shifts would be greater for the more exposed ads than for the less. Screens of lesser density might be necessary to create a spread of identification levels in a sample of the general public.

FIGURE 2. AD IDENTIFICATION AT FOUR LEVELS OF EXPOSURE




The next step in developing the CAAT will be its application to a sample of reading consumers exposed to advertisements under normal circumstances, whose recognition is measured several times after exposure, and whose claimed recognition is checked against the CAAT as an independent criterion. We would expect these criterion recognition scores to fall somewhere between Starch noted (because the CAAT has sharply reduced or eliminated non-recognition factors) and Gallup & Robinson proved name registration scores (by providing respondents with a view, though restricted, of the ad). These scores would be expected to hold up for periods of two or more weeks in the manner characteristic of recognition measures.

With fewer screens, the device would provide increasingly structured cues to the recall of headlines and other ad elements. The CAAT could also measure the visual dominance of various elements of the illustration—what comes through at which screening levels? For example, if getting the focal point of the illustration across quickly were the advertisement's objective, the investigator might find it useful to discover the levels at which the visual element in question comes through against backgrounds of different color, using different arrangements of the visual elements, and so on.

Finally, the CAAT can help evaluate the relative importance of predispositional factors which affect print ad recognition: interest in the advertised product, nearness to purchase of the advertised brand, dissatisfaction with a presently used brand other than the one being advertised, and so on. For example, could respondents close to purchase of a brand correctly identify its ads more readily than respondents dissatisfied with their present brands of the advertised product class?

These additional capabilities remain to be investigated. The CAAT's sensitivity to exposure suggests that these investigations would be worthwhile.

## REFERENCES

APPEL, VALENTINE AND MILTON L. BLUM. Ad Recognition and Respondent Set. *Journal of Advertising Research*, Vol. 1, No. 4, June 1961, pp. 13-21.

LUCAS, DARRELL B. The ABC's of ARF's PARM. *Journal of Marketing*, Vol. 25, No. 1, July 1960, pp. 9-20.

# Measuring Advertisement Recognition

D. MORGAN NEU
Daniel Starch and Staff

**The recognition method of measuring ad readership has recently met with heavy criticism. The author draws on his extensive experience with the technique to rebut its critics.**

FOR THE PAST TWO OR THREE decades, two types of research techniques have been used extensively to determine readership of advertising. These methods are called recognition and recall. Practically all readership studies have been based on one or the other technique or a combination of both. Even though the uses of the two techniques have been varied and many, two schools of thought have grown up as to which is better, and many arguments have been put forth as to the advantages and disadvantages of each. Probably the most widely used readership service is the Starch Advertisement Readership Service, which uses the recognition method.

One of the advantages of the recognition method as Starch uses it is that in the survey interview the complete advertisement is observed in its true setting and situation. That is, the issue reader sees the actual advertisement in the same situation as when he went through the issue on his own volition. All the natural cues and clues are available to the respondent to help jog his memory as to whether he actually saw that particular advertisement in that particular issue of that particular publication.

Readership recognition studies as practiced by Starch are made continuously and consistently so that trends and comparisons can be made. Its extensive acceptance and use occasionally presents problems because not all advertisements will be read equally; some are obviously going to obtain higher readership results than others. Sometimes attempts to explain low readership scores result in criticism of the research technique itself. Some of the areas of criticism that have been raised in recent years are discussed in this article.

One of the approaches to discrediting the recognition technique is to try to show that some people will say they read advertisements they could not possibly have seen. Generally the technique used is to slip bogus ads in the issue of the magazine that the respondent has read and to ask him

**D. MORGAN NEU is a vice president of Daniel Starch and Staff where he directs two copy research programs, the Reader Impression Studies and the Viewer Impression Studies, and edits the house publication, *Tested Copy*. Before joining Starch in 1950, he was a research fellow of the Instructional Film Research Program at Pennsylvania State University, consultant to John Wiley & Sons, testing supervisor at the Psychological Corporation and assistant instructor in psychology at Indiana University. There he received a B.A. in 1945 and an M.A. in 1947, and in 1950 a Ph.D. from Penn State, all in psychology. He is a member of the American Psychological Association and the American Marketing Association.**

whether he has read the ad. If some respondents say they have read the bogus ads, the charge is made that the recognition technique is invalid and does not measure what it is supposed to measure, namely, the extent to which people have seen and read the ads.

The primary purpose of this article is to show that in a bogus ad study, a proper research design must take into account operating factors other than the bogus ads inserted in the magazine. One type of properly designed readership study is to present the respondent with a stimulus situation which he has experienced before and to determine from him what he remembers about that stimulus situation. Of the ads that he could have seen, which does he remember seeing? This is the type of design used by Daniel Starch and Staff.

Another type of properly designed readership study is to present the respondent with a stimulus situation, part of which he could have experienced and part of which he could not possibly have experienced. The respondent must be made aware of this situation. Otherwise we do not have a proper research procedure for measuring readership. Such a study can show us the extent to which it is possible to trick people by asking a leading question, but it does not measure the extent to which they remember seeing and reading advertisements. That is to say, a bogus ad test may tell something about the bogus ad test but it does not and cannot tell anything about how many persons read an actual advertisement when it does appear. Specifically, the scores of the bogus ads may not be "subtracted" from the scores of the real advertisements to yield a "true" readership score.

The research literature contains many articles reporting erroneous conclusions that can result from asking leading questions. Starch (1946) has reported a number of studies conducted by other researchers as well as himself in which it is clearly indicated that confusion, bias and other faulty procedures can lead to invalid results. One of these studies was a test of the extent of erroneous recognition when bogus ads are introduced. The test results clearly show that erroneous recognition of bogus ads is small when correct interviewing procedures are used. Faulty procedures which confuse and bias the respondent can increase the error three or four times.

The research studies reported in the present article are part of a series being conducted to bring up to date the evidence that Starch found in past studies.

## METHOD

### *Research Design*

The study was made to compare the difference in erroneous recognition when the respondents were told that some ads could not have been seen previously by them and when the respondents were given no introductory remarks.

A sample of 318 readers (160 men and 158 women) of the April 21, 1961 issue of *Life* magazine was used. The sample was divided into two groups; one was made aware that there were bogus ads present, the other was not.

The sample selection followed the regular Starch Readership Survey procedure. A national sample of readers of the magazine was used, the distribution paralleling the circulation breakdown of *Life* magazine by geographic areas as well as urban and suburban areas.

It was desirable to have a test magazine that appeared unmutilated and in normal condition and at the same time contain some ads that readers could not have seen. *Life* is published in several editions, each distributed in a different area. Most advertisements appear in all editions, but a few appear only in some. It was decided, therefore, to use one of the West Coast editions of *Life* in all areas of the U.S. In this manner the advertisements that appeared in this West Coast edition but not in one or more other editions were "bogus" advertisements for the sections of the country where they did not appear. As we have said, this procedure avoids the use of dummy editions or portfolios in which advertisements are artificially assembled. Furthermore, it is possible to conduct the interviewing with the regular Starch staff working under conditions identical with those existing for the program on which the regular Starch reports are based.

### *Questionnaire*

A person qualified for an interview when he was able to state definitely that he had seen or read the issue placed before him. The respondent was asked:

Have you glanced through or read any part of this particular issue of *Life*?

Once the person qualified as a reader, Group A was given the following instructions:

The copy of *Life* I have here is actually from another part of the country. It contains some advertisements which are not in the edition you read and hence could not have been seen by you. I'd like to know about the advertisements you definitely remember having seen in the copy of the April 21st issue of *Life* which you read.

From this point on, the regular Starch readership interview was conducted. Group B, the other half

of the sample, was given no introductory instructions about the advertisements but proceeded immediately to the regular readership interview.

It might be well to describe the regular Starch readership interviewing procedure because an understanding of the procedure will help clarify why recognition readership surveys conducted by organizations other than Starch may not get the same over-all level of readership as Starch.

The starting point for each respondent is determined by a rigidly controlled system of selection. The reader is exposed to each ad page and asked,

Did you see or read any part of this advertisement?

If the respondent answers "no" or is uncertain that he saw the ad in this particular issue, the interviewer must record the answer as "not read." If the respondent answers "yes" to any given advertisement, he is questioned on each component part of that advertisement. The checking of the component parts is another advantage of the recognition method and it serves several purposes. In addition to providing information for determining the extent of reading of the ad, it acts as a probe to make certain that the respondent actually saw or read some part of the advertisement in question and was not mistaken in his initial response.

## RESULTS

The research plan employed here permitted us to compare the results of the two halves of the test sample when all conditions were the same except for the introductory instructions. It also permitted a comparison with the results of the regular readership survey on the issue.

The results in Table 1 substantiate our hypothesis that making the respondent aware of the situation with which he is faced, and not leading him to believe that he could have seen or read the ads in question, will keep erroneous recognition to a minimum. The table also shows, as others have found, that poorly designed research which introduces confusion and bias can increase the error three to four times.

TABLE 1
NOTING OF BOGUS AND REAL ADS

| | | *Mean Noted Scores* | |
|---|---|---|---|
| *Group* | *Number* | *9 Bogus Ads* | *62 Real Ads* |
| Knew some ads were bogus | 152 | 6 | 27 |
| Did not know some ads were bogus | 146 | 24 | 30 |
| Regular Starch readership survey | 318 | — | 26 |

To test the reliability of our findings, a second test was made in which respondents were told that they could not have seen all of the ads in the test issue. The sample consisted of 132 men and women readers of the June 16 issue of *Life* magazine. One of the West Coast editions was used nationally for the test and the introductory remarks were the same as those used for Group A in the first study. The results in Table 2 were almost exactly the same as the results in the first test, which reaffirms our conclusion that correct instruction to the respondents will keep erroneous recognition to a minimum.

TABLE 2
MEAN NOTING OF BOGUS AND REAL ADS

| *Group* | *10 Bogus Ads* | *56 Real Ads* |
|---|---|---|
| Knew some ads were bogus (132) | 6 | 26 |
| Regular Starch readership survey (316) | — | 27 |

These studies bring the evidence up to date and confirm the results of the studies made by Dr. Starch some years ago. We must emphasize again the importance of taking reasonable care to make certain that the respondent is confronted with the same questioning situation that the researcher will later claim was the case in the research report.

## OTHER FACTORS

Other influencing factors that have been sources of controversy are respondent set, interviewer set and memory.

### *Respondent Set and Interviewer Set*

"Respondent set" is usually produced by influences present in the interviewing situation that bias and contaminate the respondent. The careful researcher will make certain that in addition to clarity of procedure and instructions to avoid bias and confusion on the part of the respondents, care is taken in the selection, training and supervision of interviewers to avoid "interviewer set." Many studies are made without checking to see whether the interviewers are qualified to do the type of study assigned to them.

For many years Daniel Starch and Staff has used a rigid system of quality control of interviewers. The details of this system were explained in the periodical, *Tested Copy* (Starch, 1957). Because of the tremendous volume of continuous interviewing done by Starch, the quality control program is probably the largest and most intensive interviewer control ever developed. We have found, for example, that it takes very thorough and careful training to prevent "interviewer set." This is a condition which means that an interviewer has a tendency to obtain a certain type of answer.

Experience in conducting a certain type of interview, such as a readership survey, is most important. We have found that interviewers conducting a readership interview for the first time will show a "set" toward getting more of one kind of response than another. Many inexperienced interviewers will find more people than it is reasonable to assume, who reportedly read most advertisements. On the other hand, a few untrained interviewers will find hardly anyone who read the advertisements.

One approach to the problem has been to study the patterns of response of interviewers. The average readership level attained for all advertisements by an interviewer on a study is compared with the average obtained for all interviewers in the same study. Interviewers with consistently extreme responses or with readership scores consistently confined within too narrow a range are thereby identified. Such interviewers must be re-trained until they achieve an acceptable pattern of response or are dropped from the staff.

A good interviewer is one whose work has a wide range of responses from low to high such as normally is the case among the readers of a publication. In other words, the kinds of responses a good interviewer will obtain is unpredictable. When we can predict the kinds of responses an interviewer will get, she is revealing a "set" and is not conducting the interview properly.

The evidence from our quality control studies leads us to believe that organizations that infrequently conduct readership studies do not avoid the problem of "high or low" readership interviewers. When the number of interviewers used in a study is small and differences between the answers reported by interviewers are great, misleading conclusions may result. If we assume there are certain individuals who have a tendency to claim recognition of magazine advertisements, and there are others who have a tendency to deny having seen such advertisements, regardless of whether they were actually exposed to them, one would expect such a "respondent set" to be evident equally in the work of all interviewers. The results for the entire sample of an experimental study (Tables 3 and 4) would appear to bear out the conclusion that there is a strong tendency for respondents who erroneously report recognition of many bogus ads to also report recognition of more real ads. However, if Table 4 is broken down to show the distribution of respondents who were questioned by interviewers

TABLE 3

NOTED READERSHIP OF REAL ADS BY RESPONDENTS CLASSIFIED AS TO READERSHIP OF BOGUS ADS

| *Bogus Ads Noted (of 10)* | *Number of Respondents* | *Per Cent of Real Ads Noted* |
|---|---|---|
| 3 through 10 | 45 | 39.0% |
| Less than 3 | 71 | 23.2 |

TABLE 4

DISTRIBUTION OF RESPONDENTS BY BOGUS AND REAL AD READERSHIP

| *Bogus Ads Noted* | *Real Ads Noted: 30% and over* | *Real Ads Noted: Less than 30%* | *Totals* |
|---|---|---|---|
| *All Interviewers* | | | |
| 3 through 10 | 28 | 17 | 45 |
| Less than 3 | 13 | 58 | 71 |
| Total | 41 | 75 | 116 |

TABLE 5

DISTRIBUTION OF RESPONDENTS BY INTERVIEWER READERSHIP LEVEL AND HIGH AND LOW BOGUS AND REAL AD READERSHIP

| *Bogus Ads Noted* | *Real Ads Noted: 30% and over* | *Real Ads Noted: Less than 30%* | *Totals* |
|---|---|---|---|
| *High Interviewers* | | | |
| 3 through 10 | 10 | — | 10 |
| Less than 3 | — | — | — |
| Total | 10 | — | 10 |
| *Medium Interviewers* | | | |
| 3 through 10 | 18 | 13 | 31 |
| Less than 3 | 12 | 33 | 45 |
| Total | 30 | 46 | 76 |
| *Low Interviewers* | | | |
| 3 through 10 | — | 4 | 4 |
| Less than 3 | 1 | 25 | 26 |
| Total | 1 | 29 | 30 |

classified on the basis of previous performance as "high," "medium" or "low" readership interviewers, it is evident as shown in Table 5 that another very important factor, "interviewer set," is operating.

All respondents, for example, who were interviewed by high readership interviewers were high on both bogus and real ad readership. On the other extreme, 25 out of 30 of the respondents who were interviewed by low readership interviewers were low on both bogus and real ad readership. Still a third pattern of response was evidenced by the respondents who were interviewed by medium readership interviewers. It is clear from this example what an important influence high and low interviewers can have on readership results.

In summary, when the results for all interviewers are observed, there appears to be evidence for a strong "respondent set." However, when interviewer differences are taken into consideration, the results are shown to be produced in large measure by a still stronger "interviewer set." It is also abundantly clear from such examples that the greater the number of interviews that are conducted by inexperienced interviewers with a rigid "interviewer set," the greater are the chances for misleading results in readership studies.

### *Memory*

In 1956 the Advertising Research Foundation sponsored the *Printed Advertising Rating Methods Study* (PARM) which was designed to duplicate the methods of several commercial readership services and to check their consistency. Some major findings of the PARM Study, as reported either by Sherwood Dodge (1956), Chairman of the PARM Committee or by Dr. Darrell B. Lucas (1960), then Technical Director of the Advertising Research Foundation and a member of the PARM Committee, are as follows:

1. The sampling procedure used by Starch is satisfactory.
2. The Starch method of qualifying issue readers is adequate for the purpose of obtaining recognition ratings of this type.
3. Starch is utilizing the time period suitable to its own research method.
4. The Starch method is reproducible.
5. It defines its methods and procedures clearly.
6. It is practical for day-to-day operation.

In general, to quote Lucas (1960, p. 15): "The correspondence between PARM and the regular Starch ratings was generally accepted as an endorsement of this commercial service."

One finding of the PARM Study has produced some discussion, however. The PARM tabulations showed the lengthening time span since last reading took place brought no slump in average ratings. Since this finding is contrary to the generally accepted theory that memory should decline with the passage of time, the PARM analysts concluded that there was more likely "something in the interview situation itself which produced the erratic behavior of recognition scores" (ARF, PARM, vol. 4, Bulletin No. 6, May 1957, p. 3).

In the past four years Starch has made a number of checks on this point. The results of a few of these studies were published in the January 1958 issue of *Tested Copy*. None of the Starch checks show the results obtained by PARM. All of the Starch studies show a lowering of the readership scores with the passage of time. We, therefore, would conclude that something in the PARM interviewing situation produced the "erratic" behavior. However, it appears the discussants of this memory problem such as Lucas (1960), Appel and Blum (1961) and Simmons (1961) choose to ignore the Starch findings and continue to speculate about the PARM results.

Regardless of which results are correct, Starch or PARM, we agree with Lucas that there is a more fundamental question involved with the memory factor. Lucas states (1960, p. 15), "Recognition is a type of memory, and memory should decline with the passage of time." Simmons says (1961, p. 83), "In the July 1960 issue of the *Journal of Marketing*, Dr. Lucas called attention to the fact that the traditional psychological laws of forgetting, reflected by the Ebbinghaus curve of forgetting, should be obeyed by a correct measure of memory of an ad resulting from exposure." Lucas did not say exactly this in his article, but for purposes of discussion let us assume that basically Simmons and Lucas are talking about the same thing.

Much has been writtten about so-called laws of forgetting and retention both inside and outside the field of psychology. Most of it has been built up from hearsay and half-truths with little or no basis in fact. Terms like "traditional," "classical," "principle," "laws" are thrown around in a careless manner. Certainly a statistician of Simmons' reputation would not consider anything a principle, much less a law, based upon a sample of one. But it is a fact that Ebbinghaus constructed his "famous" curve from one case and this case was himself. Moreover he used nonsense syllables as the subject matter. What stimulus material could be farther away from advertising copy than nonsense syllables? This might be a good spot for a facetious remark. The truth is that there is absolutely no reason to expect a forgetting curve based on nonsense syllables to have any relationship whatsoever to a forgetting curve based on advertising copy. For the latter is carefully and deliberately designed to contain the most potent attention-getting, interesting and meaningful material that advertising communication experts can devise.

Ebbinghaus himself (1913) was dismayed at the attention given to his little experiment. He was well aware of the limitations of his method and results, and warned against too wide a generalizing from them.

One of our most important experimental psychologists, E. L. Thorndike, after reviewing and duplicating the work of Ebbinghaus and many others in this field came to the conclusion that ". . . the only facts they display with any unanimity and brilliance are the apparent complexity and variability of deterioration of memory by disuse and our lack of knowledge about it" (Thorndike, 1913, p. 324). He concluded further that no one curve of forgetting is applicable to different functions or to the same function at different stages of advancement in the learning process. The psychological literature is full of skepticism and criticism of the indiscriminate adherence to curves of forgetting or retention.

It is interesting to note that only this year a writer, James Playsted Wood, published a book for which he did considerable research of the literature on memory and forgetting. He concludes (1961, p. 28): "Because of its attractive clarity and definiteness, the Ebbinghaus Curve and Law, and the comparable findings announced by his successors and imitators stand, for too many, as the law of the Medes and Persians. They have been taken for undeviating fact applicable to many kinds of material under varying conditions and in all situations wherein memory is involved."

He further remarks: "Studies of advertising effectiveness have generally remained with the Ebbinghaus doctrine and methodology because their arithmetical findings lend themselves more easily to textbook explanations or the uses of promotional research, and because there is as yet no vocabulary in which the larger aspects and implications of memory can be patly phrased. These cannot be expressed in a scientific-looking curve, proclaimed in a law or pontificated into a pie-chart." (Wood, 1961, p. 58.)

Certainly, it is to be expected that forgetting can take place for all things, but the fact is that forgetting is found to vary with the character of the event, the character of the individuals involved, the conditions between the stimulus and remembering situations, and the conditions at the time of measuring retention.

Since all these events are operating in ad recognition and advertising can be a very potent stimulus, it is possible that ad remembrance could drop off, stay the same or even increase.

Ad remembrance cannot be taken as a mechanical thing. The test of it cannot be a mechanical test. The physical presence of the same ad before two individuals does not mean that it is the same ad for both individuals. This is a hard concept to grasp, especially for those not trained in the dynamics of behavior. Lucas says (1960, p. 16), "It may seem reasonable to recommend that recognition ratings be used more as an indicator of *behavior* than *recognition*." We must remind him that recognition is behavior.

Lucas goes on to say that perhaps recognition ratings are a guide to some kind of psychological contact more substantial than mere visual exposure. We submit that this is exactly what the Starch readership scores are as obtained by the Starch recognition method. Starch is not measuring the passage of eyes across the page. Mere "visual exposure" or the passage of eyes may mean something for promotional purposes but it means nothing in terms of advertising effectiveness (unless we accept the widely questioned doctrine of "subliminal" influences). A psychological contact—the behavioral event that is reported when a person tells what he has seen and read—is the act that must take place before the advertising message can hope to have any possible influence upon the reader. And as we have tried to bring out in this article, the selection, training and supervision of interviewers, reasonable emphasis on respondents' certainty, and checking of component parts, are of prime importance in readership studies and keep erroneous or confused recognition to a minimum.

## REFERENCES


ADVERTISING RESEARCH FOUNDATION. *A Study of Printed Advertising Rating Methods.* New York: Advertising Research Foundation, 1956.

APPEL, VALENTINE AND MILTON L. BLUM. Ad Recognition and Respondent Set. *Journal of Advertising Research,* Vol. 1, No. 4, June 1961, pp. 13-21.

DODGE, SHERWOOD. *Facts and Fiction About Ad Ratings.* Unpublished speech to the American Association of Advertising Agencies, White Sulphur Springs, West Virginia, April 27, 1956.

DODGE, SHERWOOD AND WILLARD R. SIMMONS. Dodge & Simmons Debate Recognition and Exposure. New York: *Media/Scope,* May 1961, pp. 50-51, 82-86.

EBBINGHAUS, HERMANN. *A Contribution to Experimental Psychology.* New York: Teachers College, Columbia University, 1913.

LUCAS, DARRELL B. The ABC's of ARF's PARM. *Journal of Marketing,* Vol. 25, No. 1, July 1960, pp. 1-20.

STARCH, DANIEL. *Factors in Readership Measurements.* New York: Daniel Starch and Staff, 1946.

STARCH, DANIEL AND STAFF. Starch Quality Control Program. *Tested Copy,* No. 83, October 1957.

STARCH, DANIEL AND STAFF. Differences in Readership by Time Spent Reading. *Tested Copy,* No. 84, January 1958.

THORNDIKE, EDWARD L. *Educational Psychology, Vol. 2, Psychology of Learning.* New York: Teachers College, Columbia University, 1913.

WOOD, JAMES PLAYSTED. *Advertising and the Soul's Belly.* Athens, Georgia: University of Georgia Press, 1961.

# Recognition of Ad Elements: Recall or Projection?[1]

ERIC MARDER and MORT DAVID
Eric Marder Associates Kenyon & Eckhardt, Inc.

**In this eight-city study of the parts of an advertisement, the authors find most claimed readership to be a product of the respondent's imagination.**

A BASIC ADVERTISING RESEARCH PROBLEM is that of finding out which elements of an ad were actually "received" by a reader; which illustrations he looked at, which headlines and copy units he read.

Direct observation presents many problems. Even with elaborate equipment, it is difficult if not impossible to isolate the particular elements of an ad the respondent is "taking in," and laboratory methods become irrelevant if it is desired to measure exposure under normal reading conditions. Accordingly, it seems necessary to seek the desired information by asking the respondent what he has looked at or read. Various methods of questioning have been used to elicit this information. The problems involved have been widely discussed (e.g., Lucas, 1960; Appel and Blum, 1961).

First, there are methods which involve questioning the respondent about what he has looked at or read without having the ad in front of him while he is answering. These methods may be classified as "open-ended" reports. Even under optimum conditions, open-ended reports cannot serve as measures of prior behavior, but constitute measures of recall.

To overcome this difficulty, and to ascertain what the respondent actually saw when he encountered the ad, we may show him an exhibit ad and ask him which portions of the ad he previously looked

**ERIC MARDER, president of Eric Marder Associates, Inc., formerly was vice president and supervisor of market research at Kenyon & Eckhardt and before that director of theoretical research at International Research Associates. He received a bachelor's degree in electrical engineering from the City College of New York in 1946 and an M.A. in sociology from Brown University in 1950. For several years he taught physics at the University of Rhode Island and Rutgers University. He is a member of the American and World Associations for Public Opinion Research.**

**MORT DAVID is an account research supervisor at Kenyon & Eckhardt in New York. He has been with K & E for the past five years and until his recent return to New York was supervisor of research in their Los Angeles office. Mr. David received his B.A. in 1954 and M.A. in 1956 from New York University, both in social psychology, and is presently a doctoral candidate in sociology at Columbia University.**

[1] Based on a paper read at a meeting of the American Association for Public Opinion Research in San Francisco, May 19[illegible].

at or read. Measurements which employ such an exhibit may be called "aided" reports. One method of obtaining such a report is to show the respondent the ad, point to various illustrations, headlines, and copy units, and ask, "When you saw the ad before, did you look at this picture? Did you read this headline? Did you read these words?" etc.

Does this method of questioning generate a reasonably adequate measurement of what the respondent actually looked at or read, or does it measure something else entirely? If the latter, what? The purpose of this paper is to report an experiment which was designed to throw some light on this question.

## METHOD

The ad used in the experiment consisted of the following parts:

1. A large illustration of the product in the upper half of the ad.
2. A 14-word, two-line, headline in bold black type immediately below the product illustration.
3. A 14-word, one-line, sub head in smaller type below the headline.
4. Two cartoon strips (each consisting of six panels)—Cartoon I and Cartoon II, respectively—occupying the lower third of the page.
5. Small-type copy-units describing the illustrations of the cartoon strip panels immediately below the panels. (Copy I under Cartoon Strip I, Copy II under Cartoon Strip II.)
6. Three lines of small-type copy (Copy III) at the bottom of the ad.

The experiment is based on data obtained from personal interviews with a sample of 1,206 respondents. Respondents were male heads of households (50 per cent) or their spouses (50 per cent) in eight cities (one large and one small city, in each of four regions of the country). Interviews were conducted in 200 neighborhoods (clusters) in these cities. The total sample was subdivided into three experimental groups, I, II and III, of 402 respondents each. In each cluster, one male respondent and one female respondent was assigned to each of the three experimental groups.

The interview in Group I consisted of the following steps:

1. Attitude questions covering a variety of products.
2. The respondent was handed a portfolio of ads and was told, "Now I'd like you to look through this folder of ads just as you would look through a magazine. If you come across any ads that interest you, read them carefully. You may take as much time as you want." The portfolio contained 14 different ads (one ad per product category) including the test ad described above.
3. When the respondent was finished looking through the portfolio, the original attitude questions were repeated, and a number of additional questions (unrelated to the test ad) were asked.
4. At this point, the respondent was shown a photographic reproduction of the test ad (slightly reduced in size) and was asked the following question: "When you looked at this ad before, did you look at this picture? Did you read this headline? Did you read these words?" etc. For purposes of distinguishing between the ad which was contained in the portfolio and the reproduction of the ad which was used in the questioning, we shall refer to the former as the Test Ad, and to the latter as the Exhibit.

Groups II and III differed from Group I in only one respect. The headline and the subhead were eliminated from the test ad in Group II, and the cartoon strips and the copy blocks were eliminated from the Test Ad in Group III. The same Exhibit (consisting of the entire ad) was employed in all three groups.

Thus Group I was questioned with the aid of an Exhibit which was an identical reproduction of the Test Ad. Group II was questioned with the aid of an Exhibit that contained a headline and a subhead which was not in the Test Ad, and Group III was questioned with the aid of an Exhibit that contained two cartoon strips and three copy blocks which were not in the Test Ad.

## HYPOTHESES

We considered two alternative hypotheses concerning what happens when a respondent is questioned in the manner outlined above.

One possibility is that the method of questioning elicits factual reports, that is, that respondents give substantially accurate reports of what they actually saw and read during the initial exposure. If this is so, we would expect that no respondents (or very few) would report having looked at or read elements which simply were not in the test ad. The amount of measurement error, moreover, should be relatively small. The respondents were exposed to the Test Ad under controlled conditions, knew that they were likely to be questioned about what was shown to them, and were actually asked to report what they had looked at or read within a few minutes of the original exposure—an interval short enough to reduce memory distortions to a minimum.

Conversely, the possibility was considered that respondents questioned by means of an Exhibit do not give factual reports of prior behavior, but rather respond projectively to the characteristics of the Exhibit itself. If this is so, it should make no difference whether an element had or had not been in the original Test Ad, and we should obtain substantially the same (projected) reports of exposure in all groups in which the same Exhibit is used in the questioning. In that case, we might say that the

respondent is in effect responding to a hypothetical question in the form, "If you were to encounter the ad which I now show you, do you think you would look at this element, read this element?" etc. For purposes of this discussion it is immaterial whether the replies obtained in response to such a projective question correlate with what the respondent might actually see or read. For the present, we are solely concerned with assessing the degree to which the questioning procedure elicits factual reports of prior exposure or projection, respectively.

While factual reporting and projection were discussed as two extreme ideal-type alternatives, we might expect that in an actual situation both mechanisms will operate to some extent—and we may define a measure of the degree of projection in the following terms:

If the normal report of exposure for an element is N (that is, the percentage of respondents reporting exposure when the Test Ad and the Exhibit are identical), and if the projected report of exposure is S (that is, the percentage of respondents reporting exposure to an element when they could not have been exposed to it because it wasn't there), then the rate of projection P will be defined as $P = \frac{S}{N}$.

## RESULTS

Table 1 shows the results of the experiment. The first three columns, headed Group I, Group II and Group III, show the percentage of respondents reporting exposure to the various elements of the ads in these groups. Projected reports are boxed. The fourth column shows the rate of projection as defined above.

TABLE 1

PER CENT OF RESPONDENTS REPORTING HAVING LOOKED AT OR READ AD ELEMENT

| | *Group I (N = 402)* | *Group II (N = 402)* | *Group III (N = 402)* | *Rate of Projection* |
|---|---|---|---|---|
| Cartoon I | 38% | 36% | [26%] | 68% |
| Cartoon II | 37 | 34 | [23] | 62 |
| Headline | 40 | [35] | 36 | 87 |
| Subhead | 13 | [11] | 10 | 85 |
| Copy I | 4 | 5 | [5] | 100+ |
| Copy II | 3 | 3 | [4] | 100+ |
| Copy III | 2 | 2 | [4] | 100+ |

On examining the table, we see that even in the case of Cartoon Strip I (which consisted of six visually prominent cartoon panels), 26 per cent of the respondents who could not have seen the cartoon report that they looked at it, compared to 38 per cent who report that they looked at it in the control group (Group I). This results in a rate of projection of 68 per cent. This projection rate may be read as follows:

If we had conducted a measurement of exposure to the cartoon strip by means of the aided technique used here, 68 per cent of the reported exposure would have been projected, that is, could have been obtained even in the absence of any exposure whatsoever. This statement is made without prejudice to the possibility that the "measurement" may also be riddled with substantial errors of under-reporting which may, under some circumstances, compensate for the heavy projection rate.

Interestingly, the rate of projection seems to increase as the relative prominence of ad elements decreases. Thus, the rate of projection was found to be between 60 per cent and 70 per cent in the case of the relatively major illustrations (Cartoon Strips I and II); between 80 per cent and 90 per cent in the case of the headline and subhead; and substantially 100 per cent in the case of Copy Blocks I, II and III.

## CONCLUSIONS

We conducted an experiment to evaluate the relative adequacy of aided reports in measuring prior exposure to various elements of an ad. The experiment was limited to one ad and to one particular method of aided questioning. The circumstances of the experiment, however, were such as to give the aided reporting technique every benefit of the doubt. Exposure occurred under controlled conditions in the presence of the interviewer and the respondents were questioned within a few minutes after exposure. Even under these conditions, the aided reporting technique led the respondents to report not so much what they had actually read or looked at, but rather to "project" in terms of the exhibit that was shown to them. Even in the case of a prominent illustration which occupied over one third of the page, the rate of projection (the number of people who could not have seen the illustration but nevertheless reported that they had looked at it) was substantial, accounting for around two thirds of normally reported exposure. In the case of small copy units, projection was total. Within the limits of this experiment, we may conclude that while the numbers generated by aided reports of exposure to ad elements may be of some interest and value as projective data, they are substantially useless as factual reports of prior exposure.

## REFERENCES


APPEL, VALENTINE AND MILTON L. BLUM. Ad Recognition and Response Set. *Journal of Advertising Research*, Vol. 1, No. 4, June 1961, pp. 13-21.

LUCAS, DARRELL B. The ABC's of ARF's PARM. *Journal of Marketing*, Vol. 25, No. 1, July 1960, pp. 9-20.

# A Pre-Test of Four Australian Paint Ads

RONALD VICKERS
Berry Currie Advertising Organization

**In this detailed case history Mr. Vickers explains how his agency considered both qualitative and quantitative factors in selecting an advertising theme.**

MUCH ADVERTISING RESEARCH measures people's ability to recall the various components of advertisements. Useful as this information is, it provides little guidance on the extent to which respondents accept the different components and thereby change their attitudes. Since we believe that there is a definite relationship between attitudes and purchase habits, a knowledge of the attitude change brought about by an advertisement is obviously the basic ingredient in any assessment of its potential effectiveness.

In this project, we sought to obtain usable information on the state of respondents' minds both before and after being subjected to certain advertisements. To achieve this result, a new technique was developed. It proved successful, and the following case study illustrates the kinds of knowledge it can provide.

**RONALD VICKERS is director in charge of marketing and research at the Berry Currie Advertising Organization, an Australian advertising agency. English by birth, Mr. Vickers went to Australia in 1952 and was for seven years managing director of a London firm of marketing consultants. His interest in advertising research stems from a six months' stint he served as a copywriter. In England Mr. Vickers lectured regularly for the London County Council on marketing, market research, advertising and sales management. A member of the Market Research Society of Australia and the Statistical Society of New South Wales, he has contributed many articles to British and Australian professional journals.**

## BACKGROUND

The problem was to test the potential effectiveness of advertisements newly prepared for the campaign of our client, Balm Paints Pty. Limited. Although we sought specific information on the effectiveness of each of the advertisements, we were also looking for some basic principles which we could subsequently apply to future advertising for this client.

Paint advertising in Australia is extremely heavy, this industry being one of the most highly competitive in the country. All major brands of paint are reasonably well known to house owners. We felt, therefore, that it would be of little help to find out what per cent of those people who were subjected to our client's advertising merely *recalled* advertisements. Nor would it help much to find out what per cent recalled the headline and what per cent recalled the claims made for paint. Instead we wanted to know the attitude of the individual to Balm and to competitive paints before and after he was subjected to each advertisement. We asked: what effect had the advertisements on his disposition to buy our client's paint instead of the competitor's?

It was obvious that the kind of qualitative material being sought would never be obtained by straightforward interviewing methods. A formal questionnaire would provide far too constrained a mental channel for the conveyance of the type of mental data we were after.

Something in the way of "depth interviews" (a lamentable misnomer) had to be done. Accordingly, arrangements were made for the research to be conducted in collaboration with the agency's consulting psychologists.

We knew that the time and money available would keep the sample size to a minimum and that there were risks to be faced in using experimental methods. But we were not seeking statistical exactitudes. We wanted guidance which would check with our previous knowledge and experience while adding a great deal to it.

Surveys conducted by the client's research division had shown that the number of people rating Balm as the best quality paint had increased over the previous three years. This information provided some general indications that advertising had been effective in upgrading people's attitudes to Balm paint. It was also well established by research that the basic copy components used in the advertising were important to the decision-making process in relation to paint.

For example, it had been established that "wide range of colors," "durability" and "ease of application" were important benefits. On the other hand, as most Australian paint companies offered these advantages in their advertising, it was obvious that these could not, on their own, bring about brand preference.

One guide for this research was established by exploratory interviews. It was shown that people subjected to paint advertisements can be roughly divided into two major groups: those who had a more or less active intention of repainting their homes, and those who had no such intention. Not surprisingly, the first group was "receptive" to paint advertisements and even willing to read a considerable amount of body copy. The latter group, however, who had no interest in the subject, responded only negatively to advertisements—by their relative acceptance or non-acceptance of the total "image," drawn usually from a fleeting glance of the advertisement.

In view of this finding, it was decided that the sample used for testing the advertisements must comprise both groups of people. The analysis of results in relation to these groups was considered to be more important than age group and socio-economic group classifications. The latter, moreover, is a largely unreal definition in Australia.

**PURPOSE**

The study set out to ascertain:

1. the extent to which the messages contained in the advertisement were communicated to the respondents;
2. the extent to which the messages were accepted by the respondents; and
3. the extent to which the advertisement influenced the respondent's "image" of the product.

**METHOD**

Advertisements were prepared, reproduced with and without body copy, and placed in loose leaf covers so that they could be shown to respondents separately or collectively. A total sample of 250 people was interviewed, 133 men and 117 women.

Information concerning painting experience and opinions on paints was sought individually and in groups, from 200 respondents. Quantitative testing of the advertisements was also carried out on these 200 respondents.

Psychologists led eight groups of about 11 people each in lengthy free discussions of the advertisements and other relevant material. Of these 87 persons, 37 were also included in the 200 respondents mentioned above. The remaining 50 participated in the group work only and no attempt was made to obtain quantitative material from them. Each of the four advertisements was tested with body copy on 50 respondents, and each was tested without body copy on 50 different respondents. Thus each respondent saw two different ads, one with copy and one without.

Before being shown an advertisement, respondents were asked to write down what they considered the main features of Dulux (Balm paint's brand name) and Revelite (a competitive brand).

Advertisements were presented to the respondents individually. They were allowed one minute to peruse it in any way they desired, no instructions or suggestions of any kind being given. They were not told why they were shown the advertisement and seemed unaware that they would later be "tested" on it.

Immediately thereafter respondents were given a task (totally unconnected with painting or with

the home) to perform, which occupied them for half to three-quarters of an hour. They were then asked to write down what they understood to be the major features of Dulux and Revelite, care being taken to see that they could not refer to their previous notes. They were asked to write down all they could recall of the advertisement which they had looked at previously, and allowed to take as long as they wished.

Next each respondent was given a different advertisement, this time with no copy. *The respondent was asked to write appropriate wording (or body copy) for the advertisement.* This procedure was designed to find out the extent to which respondents accepted the given message of the advertisement as conveyed in the illustration and headline copy, and elaborated and exemplified it in the body copy they provided. The procedure also sought to learn how much of his own views about important features of paint the respondent projected into the advertisements.

So far as the writer knows, the idea of getting respondents to write copy for an advertisement by providing the skeleton advertisement as a stimulus, is (or was at that time) a new technique in advertising research. The essential advantage of this method is that it forces the respondent to reveal exactly how much of the advertisement's main message he has really accepted.

## RESULTS

Table 1 summarizes changes of opinion of Dulux *after* exposure to advertisements for 60 seconds.

TABLE 1
PER CENT OF SAMPLE WHO CHANGED OPINIONS

| | *Changed in Line with* | | |
|---|---|---|---|
| *Advertisement* | *Main Message* | *Details in Body Copy* | *No Change* |
| 1 | 12% | 32% | 56% |
| 2 | 52 | 12 | 36 |
| 3 | 8 | 28 | 64 |
| 4 | 16 | 28 | 56 |

Each respondent's recall report was rated on a five point scale on recall of message through the illustration, recall of message through the headlines, recall of important body copy, and general understanding of the advertisement. Perfect recall was rated at five and absence of recall at zero. Table 2 shows the average rating on each factor for each advertisement after an interval of 40 minutes.

TABLE 2
AVERAGE RATINGS OF RESPONDENTS' RECALL

| *Advertisement* | *Illustration* | *Headline* | *Body Copy* | *General* |
|---|---|---|---|---|
| 1 | 2.7 | 3.1 | 3.8 | 2.6 |
| 2 | 3.1 | 4.6 | 2.9 | 4.1 |
| 3 | 2.0 | 2.1 | 3.4 | 3.8 |
| 4 | 1.3 | 2.0 | 3.3 | 2.1 |

Table 3 shows what proportion of respondents accepted the message of the advertisement when invited to write body copy.

TABLE 3
PER CENT WHO ACCEPTED MESSAGE AND LIKED EACH AD

| | | *Liking* | | | |
|---|---|---|---|---|---|
| *Advertisement* | *Accepted Message* | *Most* | *Second* | *Third* | *Least* |
| 1 | 56% | 25.0% | 18.0% | 18.0% | 39.0% |
| 2 | 84 | 35.0 | 29.0 | 27.0 | 9.0 |
| 3 | 44 | 41.0 | 12.5 | 14.5 | 32.0 |
| 4 | 38 | 19.5 | 23.0 | 22.0 | 35.5 |

***Advertisement 1***

This advertisement, more than any of the other three, was differentially received by groups of respondents. It was ranked as the "liked most" advertisement by 25 per cent and as the one "liked least" by 39 per cent of the respondents.

Analysis of responses showed that the factor which seemed to determine whether this advertisement was "liked" or "disliked" was whether or not the reader was actively concerned with painting at

the time. Of the respondents who ranked this as the "most liked" advertisement, 66 per cent of them were either engaged in painting or were purchasing paint with the intention of painting. Of those who ranked it as the "least liked" advertisement, only 27 per cent were so engaged. This was confirmed by the results of both the recall testing and in the acceptance test of the advertisement's effectiveness in changing qualities attributed to Dulux.

Table 2 shows that respondents recalled *body copy* on this advertisement better than on any of the three other advertisements. This figure is made more meaningful when it is noted that of the respondents tested on recall of this advertisement, 48 per cent of them were involved in painting, and that their recall of body copy was much higher (4.6) than that of the not-involved group (3.0).

In terms of producing a change in readers' opinions of Dulux, it is noted that 56 per cent of the respondents tested on the advertisement showed no change, while 44 per cent expressed a change. Of these, 32 per cent attributed features or qualities to Dulux which were conveyed in the body copy. This was the highest figure for any of the four advertisements; again body copy in this advertisement was more effective than in any other ad—but only for involved readers.

When respondents were asked to write copy for this advertisement, 56 per cent of them elaborated the headline theme of the advertisement. This meant that despite the strong appeal of the detail of the advertisement for some readers there was nevertheless a fairly strong acceptance of the main message of the advertisement, "Painting Finished in Record Time."

An important point to be noted is that the degree of acceptance by the respondents of an advertisement is reflected in their preparedness to accept and elaborate the given message of the advertisement in the "copy" they write. Their refusal to mention the given message may be taken to indicate the ad's relative failure to appeal to and be accepted by the respondent. In this context, note that Advertisement 1 was second only to Advertisement 2 in acceptability. The comic strip feature of the advertisement was approved by both the involved and the not-involved groups of respondents, although the frequent criticism expressed was that too many different features were being presented.

In summary, this advertisement appealed by virtue of its comic strip appearance. This feature, containing information on paint and painting, seemed to appeal to those who were painting or about to paint. Of the four advertisements it was found to be the most effective one with the involved group.

***Advertisement 2***

EVEN IF YOUR WIFE WANTS 'IMPOSSIBLE' COLOURS
they're easy to get simply by
TUBE TINTING
DULUX PAINTS
DULUX PAINTS FOR COLOUR AND QUALITY

Of the 200 respondents, 35 per cent selected it as the most liked, and only nine per cent rated it as least liked, making it best liked of the four. On most of the other tests this advertisement also achieved high rating. We are well aware that the advertisements people like best are by no means always the most effective. However, liking an advertisement does indicate that it has some quality which attracts people and gives a favorable impression of the product.

The absence of rejection or negative response as indicated by the very few who ranked it least liked was confirmed when respondents were asked to "write" copy for the advertisement. On this test 84 per cent of the respondents wrote copy which directly reflected the given message of the advertisement as conveyed in the illustration and headlines, a higher per cent than that achieved by any of the other advertisements and more than twice that achieved by Advertisement 4.

In the group discussions, it was brought out that color was a crucial factor determining the choice of paint, especially for interior work, and that a brand which offers an extensive range of color was highly regarded. This advertisement—as the only

one featuring color—therefore had an advantage over the others, but many creative factors must also have been conducive to its success.

In recall this advertisement gained the highest average ratings for recall of message by illustration, recall by headline, and for general understanding. Only on recall of body copy did the advertisement rank lower than any of the other three. Together these results indicate that the advertisement successfully conveyed its message.

The effects of this advertisement on respondents' opinions were consistent with the results above. Table 1 shows that only 36 per cent of respondents failed to revise their opinion of Dulux after looking at this advertisement for 60 seconds. When asked to write down the qualities of Dulux paint 40 minutes after looking at the advertisement, 64 per cent gave revised accounts and 52 per cent gave in their revised accounts specific reference to tube tinting and the wide color choice available. No other advertisement achieved such a strong measure of acceptance. We believe Advertisement 2 achieved much of its effectiveness by successfully tapping two important psychological factors involved in the home painting situation. First, it allowed men to interest themselves in the detail question of color by projecting the interest in color to the wife. While it is commonly maintained that the wife chooses the colors to be used on the walls, the husband plays a larger role in color choice than is commonly believed. It seems that the husband seeks to efface his interest and influence in color choice by maintaining that the wife has made it. The advertisement, both in illustration and headline, picks up this interest in color and allows for both a conventional masculine and feminine expression of it.

Second, the advertisement had a strong appeal for most readers because it captured some of the exasperation that is so frequently experienced by the home painter. The human element in the advertisement played an important role in the advertisement's success and in all groups there was strong evidence of how apposite the emotions portrayed were, and how sensitive to them respondents were.

This advertisement achieved high ratings on all factors; from both the quantitative and the qualitative evidence, we concluded that it was the most effective of the four advertisements.

Advertisement 3 gained the highest ranking of the four in respondent liking: 41 per cent chose it as "liked most" and 32 per cent rated it as the advertisement they "liked least." The result here was comparable to that on Advertisement 1, which was both liked and disliked by a high percentage of respondents.

*Advertisement 3*

Evidence from the other tests did not support the rating mentioned above. It became apparent that one technical feature of the advertisement was partly responsible for its high rate of selection by respondents.

On the recall test this advertisement did not rate as well as some others. Average recall by illustration was relatively low as was recall by headlines. This is surprising because illustration and headline in this advertisement were physically more dominant than in any of the other three advertisements. The recall of body copy in this advertisement rated 3.4, second only to Advertisement 1, which rated 3.8. A relatively high rating of 3.8 for over-all understanding was achieved by the advertisement.

It is interesting to note how unsuccessful the respondents were in receiving the point "You'll be happy with Dulux . . . Australia's easiest house paint." Instead they tended to recall details from the small illustrations. This tendency was also noted in the test of opinion change. Of the 36 per cent of respondents whose opinion of Dulux changed after looking at this advertisement, most of them changed in response to material contained in the body copy,

whereas a change brought about by the main message was detected in four respondents only. This ad achieved the poorest result in producing change in respondents' opinions of Dulux: 64 per cent failed to revise their opinion or to attribute new or different features to Dulux.

In the test designed to assess the degree of acceptance of the advertisement and its message, this proved to be one of the less well accepted of the four. Of the 50 respondents who were asked to write body copy for the advertisement, only 44 per cent wrote copy which elaborated or exemplified the given message of the advertisement. It was as though the respondents had either not comprehended the advertisement or that they could not accept it as easily as they could the other advertisements. The copy written thus tended to reflect the respondents' own views about paint rather than the illustration and headlines.

The group discussions suggested that there was a real difficulty in accepting the message of the advertisement. This was thought to be a less personal advertisement than the other three, and though it was clear to all that the man was smiling there was a strong feeling that he lacked conviction. Such happiness in the absence of other persons was regarded as improbable; the improbability of such emotion about paint rather forced itself on the respondents.

What seemed to be even more important to the respondents was their failure to perceive the man in the advertisement as a real person. While there was strong evidence from group and individual discussions that these emotional factors appear to operate, a defect in the bromide had contributed to this result: the illustration was very much darker than it would have appeared in the press. This intensity of black may well have contributed to the advertisement being perceived as overwhelming. This same feature may explain why this advertisement was selected as the best advertisement by so many respondents when it has so many limitations. Its strong and bold features, which make it stand out more than the others, may well have been the reason why respondents selected it. (This points up the dangers inherent in using reproductions for testing, as well as the fallacy of judging advertisements merely by people's preferences.)

On all the quantitative tests this advertisement achieved the lowest ratings. The qualitative evidence suggests that not only was it the least effective of the four advertisements, but the only one which

*Advertisement 4*

evoked strong negative feelings in the respondents.

This advertisement received the lowest preference rating. Only 19.5 per cent of the 200 tested selected it as the liked most advertisement and 35.5 per cent ranked it as the one liked least.

On tests of the success of the advertisement in communicating its message as measured by the success of the respondents' recall, it proved to be much poorer than the other three advertisements. The average rating on recall by illustrations was very low (1.3) and very low on recall by headlines (2.0). It was more successful in communicating body copy (3.3) and very low (2.1), in over-all understanding. These results show that the advertisement was unsuccessful in conveying its message either because it was not understood or because of some "resistance." There was evidence to support both, but the stronger evidence was for the latter.

The advertisement was reasonably successful in bringing about a change in opinion about Dulux, but of those respondents changing their opinions, more revised their views as a result of details in the body copy than as a result of the main message. For example, more made reference to the "195 color range" than to the notion that "Dulux is glossy as tiles and stays as clean."

The negative attitude to the advertisement most clearly shows in the respondents' own copy for the advertisement. Only 38 per cent wrote copy which

elaborated or exemplified the main message of the illustration or headlines. Strong negative responses to this advertisement were encountered in all of the groups and similar expressions were detected in individuals when they were being tested. Women in particular often reacted to the advertisement by scoffing at the girl portrayed, by maintaining that she was a "glamour puss," wouldn't know how to hold a paint brush and that no one would wear beads like that or put on a black dress for painting.

Only rarely did a respondent see the girl in the advertisement in the perspective that was intended —a modern young housewife "appreciating" the glossy tile-like quality of Dulux. Repeatedly the girl was made to be a home painter; made to be like the respondents. It was as though they could not tolerate her to be superior to themselves. She could not be immaculate, groomed and poised when they were anything but.

While this rejection of the girl and her message was most marked among women respondents it was by no means absent among men. While women tended to take the matter personally, there was a tendency among the males to deflect their feelings and to treat the girl and the message as unreal and a joke. The feelings expressed about this advertisement left no doubt that the lack of acceptance of the message of the advertisement was mainly the result of strong emotional resistance to the illustration.

## CONCLUSIONS

Not surprisingly, some of our findings confirmed advertising text book maxims. However, the results were considered extremely useful by our agency copy writers and by the client's advertising executives. Apart from the specific guidance provided in relation to revision of the advertisement tested, the research indicated a number of general principles:

1. Advertisements which secure the reader's emotional involvement are far more successful than those which do not.
2. To secure this emotional involvement, a situation—thoroughly believable in terms of real life—can be a very successful device.
3. Apart from emotional factors, the involvement of potential buyers can best be secured by portraying a human situation which might actually occur.
4. Successful involvement of the reader necessitates acceptance of the various ingredients of the advertisement and acceptance of the whole. The idea, illustration and the copy must square with the reader's own experience.
5. The emotion(s) around which the advertisement is built (e.g., happiness) must have a believable cause. The emotions must flow from the benefit received from the product, *not* from the product itself.
6. The message of the advertisement must be unambiguous and clear.
7. People "see" advertisements through a mental screen comprised of their own personality and interests. People actively interested in painting may see the advertisement quite differently from those not so interested.
8. The old adage, "anything that can be misunderstood will be misunderstood," applies to advertisements, and in particular, to headlines.
9. The fact that the advertisement is for paint—and the brand name—should be noticeable *at sight.*
10. People who read *all or most* of an advertisement, do not necessarily remember the heading or main illustration. They remember the points which are significant *to them,* even if these are given minor mention in the advertisement.
11. If realistic art is used, the illustration used must be believable and people must be "real." A wife must look like a wife and not like a mistress. (This does not mean, of course, that she should not be attractive!)
12. The expressions depicted on people's faces should be clear in themselves, and not depend on copy for explanation.
13. Small comic strip illustrations increase readership of body copy.
14. The headline may be inadvertently camouflaged by the adjoining illustration. Alternatively, its readership may be increased by the nature and juxtapositioning of the illustration.
15. Spacing, positioning and the style and size of type can be very important influences in determining the extent of readership of the headline.

*Cautions:* Our respondents had to be reasonably intelligent and well educated, which meant that the sample was not statistically representative of any universe. They were paid for participating. Skilled, highly paid people had to be employed for the interviewing work. Consequently a large sample would be expensive. These disadvantages were, we thought, greatly overshadowed by advantages. Certainly both agency and client now know more about paint advertising than either did before the research was undertaken.

# Three Ways to Improve Ad Pre-Tests

**BILL STEVENS** and **JOEL AXELROD**
New York University — Benton & Bowles, Inc.

Drs. Stevens and Axelrod advocate three simple principles which, they say, would remove much of the present confusion about the purposes and methods of copy testing.

THIS ARTICLE is about an activity almost all advertising agencies use to aid their decisions—the pre-testing of advertising. Though it has been of considerable value to management, there is much room for improvement. The contention of this paper is that three relatively simple changes would lead to such improvement.

Copy pre-testing is ordinarily requested by an account executive, copy writer or client. One of them comes to the copy tester with any of a variety of problems. In the case of a single theme or ad they want to know if it is "good." In the case of two or more ads or themes they want to know which is "better" or "best." They mean a variety of things. Will the ad "stop" people, or which ad has more "stopping" power? Do people understand it, or is there a difference in degree of comprehension? Do they like the ad? Do they like one ad more than another? Do people believe a claim, or believe one claim more than another? Does the ad interest them in the product, or does one ad create more product interest than another? Does the ad create the right impression about the product? There are many others. Which questions they consider most important vary from one situation to another.

What they are really asking, indeed what they *must* be concerned with as businessmen, is will the ad sell the product? Which ad will sell more? However, they seldom ask this question explicitly because they are aware that copy research cannot generate such information except under unusual and expensive circumstances.

In translating these questions into research procedures, the copy tester usually starts by trying to

**BILL STEVENS is an assistant professor of marketing in the Graduate School of Business Administration at New York University, where he teaches industrial marketing, marketing research, and advertising in marketing management. He was associate director of research at Young and Rubicam for eight years, and more recently held the same position at Dancer-Fitzgerald-Sample. Professor Stevens received an A.B. from the University of Kansas and a Ph.D. in social anthropology from Harvard in 1950. He is a member of the American Marketing Association and the American Association for Public Opinion Research, and is active in the A.M.A.'s Marketing Educators' Forum.**

**JOEL N. AXELROD is supervisor of advertising research at Benton & Bowles. He received a B.A. from Brown University in 1954 and a Ph.D. in psychology from the University of Rochester in 1958. While doing graduate work he was associated with the Rumrill Company and the Psychological Corporation. Before joining Benton & Bowles in 1960 he was at Dancer-Fitzgerald-Sample. Dr. Axelrod is a member of the American Psychological Association and the American Marketing Association.**

learn as much as possible about the background of the problem. What is the competitive situation, the history of the campaign, the copy platform, the marketing strategy? At best he is only partially successful, even if all parties concerned try to be helpful. There are so many prior decisions that have never been committed to writing, so many unresolved marketing decisions, and so many nuances, that often he must settle for fragments. In any case, once he has gleaned as much background information as possible, the copy tester must begin to deal with the specific mechanics of his research project.

### *What to Measure?*

He must first decide what to measure. Usually he chooses such things as playback from the ads or themes, interpretations, interest, believability, feelings of confusion, interest in or liking of the product, etc. Such measures have come, through custom and tradition, to have some degree of acceptance as valid measures of advertising effect. Seldom, however, is it specified which measures are evaluative (measures of success) and which are diagnostic (explanatory). This proliferation of measures and lack of clear discrimination among them reflects a widely held assumption that advertising works in such complex ways and along so many dimensions that many measures are needed to assess its effects.

After the questionnaire is written there remains the job of selecting the sample and putting the job into the field. Limited in time and budget, the copy tester does not even try for a projectible sample. At best he is able to obtain several hundred clustered interviews. When the interviews return, he supervises the preparation of the codes and the tabulation of the responses. Then he must interpret the results and make recommendations.

Since a multitude of questions were asked, with none being clearly defined as more important than any others, the decision as to how good an ad is, or which is better, often depends upon the question emphasized. Since the research was probably done to settle a difference of opinion, each contending party can choose the result of that question which supports his point of view. In other words, it is difficult for the copy tester clearly and unequivocally to interpret the data. Rather, the interpretation and resulting recommendations can easily become a function of the relative agency prestige of the parties involved.

A related and serious problem is that of who in fact should do the interpretation. The copy tester has the technical proficiency but may lack the marketing background. On the other hand, the research user has the marketing background but lacks the technical proficiency. Neither is satisfactory.

If the copy tester does the interpretation, the user may find that he is presented with answers which are complicated, ambiguous, irrelevant or lacking in generalization. If the user interprets the data, he may, because of his lack of technical knowledge, overgeneralize, draw incorrect inferences or use only part of the data. He may well mislead himself into making the wrong decision.

The rest of this article will be devoted to suggestions for increasing the value of copy testing to management. These suggestions are hardly revolutionary, but experience indicates that they can make a significant difference. Primarily these suggestions are in two areas: maintaining a clear philosophy of advertising, and developing a closer working relationship between the copy tester and the research user.

### *Formulate Objectives*

A fundamental step for increasing the value of pre-testing is for the practitioners to have a clear and explicit theory of advertising's objectives.

Ultimately, of course, the objective of advertising is to produce profits by increasing sales. But it is only one of a myriad of marketing factors which bring together the product and the consumer. To mention but a few others, there are promotional activities, competitive advertising, shelf position, the remarks of friends, distinctive and appropriate packaging, and even temperature. All can and do intervene between the advertising and the purchase of a product. The relationship between advertising, sales and profit is complex and indirect, and therefore not a useful criterion for routine copy testing.

So it is necessary to use the more immediate effects of advertising as the criteria by which to test copy. As previously noted, however, copy testers have tended to work under the assumption that the immediate effects of advertising are extremely complex and varied and that, therefore, there is no standard criterion or criteria which can be used to measure its effects. We contend, to the contrary, that all advertising has two direct and immediate objectives which can readily be used as criteria.

The most immediate objective of all advertising is that it be perceived and remembered; this is a necessary but not sufficient requirement. The other

objective of all advertising is to create a more favorable attitude or predisposition toward a brand.

Accepting this philosophy of advertising has several immediate advantages. The copy tester will no longer have to be concerned with *what* criteria to use in his research. Rather, he can set himself to the task of selecting the best specific instruments for measuring each criterion. Which scales or other measures will most accurately predict the memorability of an ad? Which measuring instrument is most sensitive to the effect of an ad on attitude—buying interest scales, "like" scales, semantic differential scales, psychogalvanic skin response or what? Also he can now build into the tests diagnostic measures to help explain the criteria results. No longer will he have to be concerned about their being misconstrued as criterion measures with all the resultant indecisiveness and dissent. As a result, then, of accepting clear and unambiguous criteria for the measurement of effect, the practitioner is able to present clear, concise, and relevant answers to the research user. He can indicate relative values among ideas, themes, or advertisements and offer diagnostic explanations for the results. For his part, the user can readily understand what has been measured and can therefore make better decisions.

### *Develop Perspective*

It was noted before that it is difficult for copy pre-testers to get an adequate grasp of the marketing and advertising context in which a specific problem is centered. This context is necessary for several reasons.

A very basic reason is that it is relevant to selecting an appropriate sample. If due to distribution patterns, specific product features, or budget size, a brand cannot or is not intended to compete for the great middle majority of consumers, it may be necessary to determine which ad or theme appeals to a selected group of potential consumers. In such cases, to test using the middle majority would be inappropriate. A campaign may be designed to counteract inroads made by a specific competitive brand; this would be highly relevant information. A campaign may be designed to expand the market among a given segment of the population. A clear-cut briefing on this market would be more important than large expenditures for random sampling with wide dispersion. Such briefing might also bring about certain marketing decisions which had been begged for many years.

Another equally important reason for knowing the context of the research is to help the practitioner interpret the findings. This is particularly important when the research is technically complex and must be interpreted by the researcher rather than by the user. For example, not knowing the context of the problem, the copy pre-tester would probably recommend the theme which performed best. Knowing the marketing context, knowing that a theme very similar is being used by a competitor with a much larger budget, he might recommend a theme that appealed to a specific smaller group or a group that could be reached through specific media and through specific distribution patterns. He might not recommend the superior theme because its advantage was not sufficient to compensate for the fact that a shift in the media selection and pattern of distribution would be required to make use of it. Knowing the context of the research, the copy tester can make a positive contribution to the decision-making process.

To accomplish this objective is not simple. Yet its achievement involves a relatively simple administrative agreement. The copy pre-tester has to become a functioning part of the account group, rather than a resource person who is called upon for limited technical activities. However this is accomplished, it will have two beneficial effects. The copy tester will become more of an advertising or marketing man. The user will better understand the values of copy pre-testing and may even come to think of copy pre-testing as an integral part of his work.

### *Make Assumptions Explicit*

It was noted that in carrying out copy pre-testing, assumptions are frequently made regarding the effects of execution, the relationship between practical and mood ideas, and how advertising operates over time. The difficulty is that these assumptions are implicit rather than explicit in the minds of both the copy tester and the user of the research. Making them explicit would be beneficial in that some research would not be done at all, other research would be recast, and the results of all research would be interpreted more meaningfully.

When themes, rough ads, or rough commercials are tested, the implicit assumption is that execution is no problem. Both themes are equally susceptible to good execution. The rough execution does not impair the basic clarity of one ad more than it does another. The persuasiveness of both commercials is equally impaired by the rough form. Obviously, this is not always true; perhaps it would be better to say that it is seldom true. Experienced

pre-testers know that certain ideas can be understood, felt and reacted to only when presented in fully developed advertisements. To allow a potentially successful idea to be eliminated in a theme test because it depends almost entirely for its effect on execution is something that the copy tester must warn against. Similarly, the choice of one commercial over another when the execution may have unduly impaired the performance of only one is something else the copy tester must prevent if possible.

When ads which are based on "logical," "sensible," product benefits are compared with ads which are based on mood or feeling tone, the assumption is made that respondents can, in a meaningful way, discriminate between them. Such a comparison would undoubtedly have led to the elimination of many successful, emotionally based advertising campaigns had not experienced advertising judgment intervened. This does not mean that logical ads should not be tested against mood ads. It does mean that a respondent should not be asked to choose between two such ads or themes; and that great care should be taken in interpreting the results of research involving even sophisticated designs. The copy tester has to watch for evidence that respondents may be using different frames of reference in evaluating the two types of ads or ideas; he has to look for signs that they are rejecting the mood idea in order to appear to be logical, informed, sophisticated consumers. The copy pretester should forewarn the person requesting the research when there is a possibility of this type of bias. He should point out, when such a bias does occur, that it invalidates the findings.

In varying forms, old ideas are frequently compared with new ideas. Sometimes it is a proposed campaign against a current campaign, in either theme or ad form. Often it is a proposed campaign against a competitor's long-term campaign, particularly when his sales are growing. But advertising operates differentially over time. That this fact is often overlooked is only natural since we don't know precisely how advertising operates over time. All we know is that an idea has a different effect on the nth presentation than on the first. If it is a complex or novel idea, it may not "move" people until they have been exposed to it frequently and and a chance to understand it. On the other hand, if it is a simple, clear idea, it may move people immediately. A novel or complex idea which requires a number of exposures before it begins to be effective is not necessarily a bad idea; quite the contrary may be the case. What all this means is that the practitioner has very special problems when testing new versus old ads or themes. Where such testing is unavoidable, he should make these problems explicit and be cautious in drawing conclusions from the data. One solution, though perhaps impractical in smaller organizations, is to test new approaches routinely at their inception. With a backlog of such tests, one could compare themes or ads at the same point in their development.

### *Conclusion*

Thus, it is possible to increase the value of copy testing to agency decision makers through three relatively simple steps. First: accept that philosophy of advertising which says ads have two immediate and easily measurable functions; to be perceived and remembered, and to produce a more favorable attitude or predisposition toward a product. Second: associate the copy tester more closely with the account and creative groups so that he may be more clearly aware of the context in which the research is done. Third: make explicit various assumptions regarding the effects of ad execution, practical versus mood ideas, and how advertising works over time on the outcome of copy research. The combination of these three steps should facilitate and make more efficient and fruitful pre-testing of copy.

*In our society the policy maker is dependent to an increasing extent on his subordinate's conception of the essential elements of a problem . . . Pragmatism, at least in its generally accepted forms, produces a tendency to identify a policy issue with the search for empirical data.*

—HENRY A. KISSINGER

# Consumer Income Data

INGRID C. KILDEGAARD
ARF Research Statistician

**Expensive to collect, difficult to obtain by direct questioning, income statistics are nevertheless reported thoroughly by the U. S. Government. Miss Kildegaard tells how to get this information from five of their publications.**

PEOPLE ARE MARKETS—provided they have money to spend. Advertisers know that personal income means buying power, and that shifts in buying power dictate reallocations of advertising expenditures. Such shifts occur constantly, from year to year, from place to place:

In 1960, personal income in the United States rose to 402 billion dollars, an increase of over 75 per cent in ten years.

1960 national per capita income was $2,223, 49 per cent higher than in 1950.

About ten million families reported 1959 cash incomes of less than $3,000, a drop of six million families in ten years. In a third of the ten million, the household head had reached age 65.

In 1959 the concentrated highest income groups were in the New York Standard Metropolitan Area. More New York tax returns reported incomes of $50,000 or more than in Chicago, Los Angeles and Philadelphia combined.

Mississippi's 1960 per capita income was 43 per cent below the national average while Delaware's was 36 per cent above.

Between 1949 and 1959, the average money income per family rose from $3,100 to $5,400. Adjusted for price changes, this is an increase of 42 per cent.

Federal income statistics are voluminous. Five sets of data currently available are described below. While all are valuable ammunition in setting sales quotas or planning advertising strategy, they differ in what they report and in how they define income.

*Current Population Reports: Consumer Income,* issued annually by the Bureau of the Census, give size distributions of national family and personal income. The number of consumers falling in each income layer is shown by age, sex, color, size and type of family, occupation and industry groups, urban and rural. Besides reporting distribution of income by small size groups, median incomes are estimated from a sample of the civilian household population. Each year questions on income are asked of about 26,000 respondents, age 14 or over, in the March Current Population survey. An advance report on 1960 income was issued in June; the full report is scheduled for late 1961 release.

*1960 Census of Population* provides geographic detail on income as well as distributions of the population by income size. The 1960 census income statistics are now being published in 53 reports—one for each state, the District of Columbia, Puerto Rico and a United States summary. Distributions

| Publication: | Current Population Reports: Consumer Income | 1960 Census of Population—Series PC (1) — C |
|---|---|---|
| **Issued by:** | Bureau of the Census | Bureau of the Census |
| **Source of Data:** | March Current Population Survey; about 26,000 household interviews | Population census, based on 25 per cent sample |
| **Frequency:** | Annual | Every 10 years |
| **Coverage:** | Civilian household population | Civilian total population |
| **Reporting Unit:** | Families, persons | Families, persons |
| **Definition of Income:** | Money income from: wages and salary, net receipts from self employment, and other sources including rentals, interests, pensions | Same as Current Population Reports (column one) |
| **Area Detail:** | U.S., urban and rural, four census regions | U.S., state, county, city, metro area |
| **Limitations:** | Response errors, especially those due to memory loss | Response errors due to memory loss; infrequency of reports, especially with today's shifts in population and income |

are given of both personal and family income by state, county and standard metropolitan statistical area (the information is also available by census tracts for most SMSA's). In those areas where the data permit, results are shown for their urban, rural nonfarm and rural farm parts. The same geographic detail is provided for personal income by sex, occupation groups and sources of income.

Other general social and economic characteristics of the population (education, family composition, employment, occupation, mobility, place of work and transportation to work) are also presented in this series. The entire series of reports will be out by spring 1962.

*1959 County Business Patterns* report detailed payroll data from social security tax returns by kind of business and small areas. These reports were described in the September 1961 issue of this *Journal.*

*Statistics of Income—1959—Standard Metropolitan Areas,* a new publication from the Bureau of Internal Revenue, gives distribution of income tax returns for each of the 100 largest standard metropolitan areas. For 18 income classes, the number of returns is reported as well as the aggregate totals of adjusted gross income, exemptions, taxable income and tax paid. This report provides added detail on high income groups, showing five income classes beyond the Census cut-off of "25,000 and over."

*Survey of Current Business,* a monthly publication of the Office of Business Economics, publishes data on aggregate and per capita income. The August issue reports state data by industry and by source with estimates of disposable income, savings and consumer expenditure. More detailed tables on national aggregate distributions appear in the July issue.

| 1959 County Business Patterns | Survey of Business | Statistics of Income 1959: Standard Metropolitan Areas |
|---|---|---|
| Bureau of the Census and Bureau of Old-Age and Survivors Insurance | Office of Business Economics, Department of Commerce | Bureau of Internal Revenue |
| Old-Age and Survivors Insurance records plus Census data | Component estimates based on government and business data | Income tax returns |
| Every two or three years | Every July (state) and August (national) | First report—may become annual |
| Employees of organization covered by OASDI | Total population | Population with income over $600 |
| Employees | Family (national) and personal (national and state) income | Aggregate dollars by returns |
| Taxable payroll (up to $4,800) | Money income (less social security) plus income in kind, such as food or lodging | Money income and net capital gains (less veterans payments, social security, relief, etc.) |
| U.S., state, county | U.S., state | 100 largest standard metropolitan areas |
| Earnings in excess of $4,800 not reported; includes only employees covered by social security; income other than earnings not reported | Inaccuracies and gaps in source data; skewed distributions make per capita and mean family income atypical | Returns are neither individuals nor families since some file separate and others joint returns; persons earning less than $600 not covered; some types of unearned income omitted |

***Guides for Advertisers***

Direct comparability between the sets of income estimates cannot be expected. As the chart shows, they differ in definitions, coverage, collection and reporting methods. So will the advertiser's use of the data.

For forecasting the consumer market, he needs the up-to-date trend data from the Office of Business Economics' income series. The industrial advertiser's potential customers are best described by County Business Patterns. For selecting areas to promote high-ticket goods, the high income data from the Bureau of Internal Revenue serve as a blueprint. The 1960 Census Reports contain the real meat for advertisers of consumer goods, especially if his sales are correlated with income levels. These data permit him to rank areas by sales potential, to allocate his advertising budgets and even evaluate the performance of that advertising.

More details on federal income data and their applications are in a booklet just published by the Office of Marketing Services. State and commercial sources are also described *(Personal Income: A Key to Small-Area Market Analysis).* This publication, as well as all others mentioned, may be ordered from the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C.

# Analysis of Variance

GWYN COLLINS
ARF Research Mathematician

**Any result we prize usually has many causes. Their contributions can be isolated by a statistical tool which is based on the simple fact that we can add sums of squares.**

ANALYSIS of variance is a technique for learning how important each of several factors may be in producing a result. Its fundamental idea is altogether simple, its development quite recent and its theoretical foundation a single mathematical theorem. It has wide application; regression and correlation, for instance, can be regarded as aspects of it. It is the basis of all modern experimental design.

The variance of a set of numbers is a measure of the way those numbers are scattered, a set of numbers which are widely scattered having a higher variance than one which is closely grouped together. The term has been given a precise meaning which enables the variance to be calculated: it is the sum of the squared differences between each number and the average of the set, divided by the number of independent members in the set.

Thus the variance of the five numbers 1, 2, 3, 4, 5 is one fifth of the sum $(3\text{-}1)^2 + (3\text{-}2)^2 + (3\text{-}3)^2 + (3\text{-}4)^2 + (3\text{-}5)^2$ which is ten fifths or two.

Similarly the variance of 7, 9, 11, 13, 15 is one fifth of $(11\text{-}7)^2 + (11\text{-}9)^2 + (11\text{-}11)^2 + (11\text{-}13)^2 + (11\text{-}15)^2$ which is 40 fifths, or eight. This larger variance reflects the greater scatter of the numbers 7, 9, 11, 13, 15.

The variance, of course, is not the only measure of scatter. But it has many convenient properties which have given it a special place in the theory of statistics.

One particularly useful quality, and the one used in analysis of variance, is associated with the sums of the squared deviation from the mean before they are averaged to give the variance. In fact the analysis of variance could be more accurately, though less elegantly, described as the analysis of sums of squares. For these sums of squares are "additive"—a description which will be illustrated extensively. This term, "the sum of squares," will crop up with great regularity in this article. There is no statistical term for it, though "deviance" and "squariance" have been advocated; here we shall simply refer to it as the sum of squares and it will usually be quite clear which squares are being summed.

In the examples above, the sum of the squared differences between each number and the average of the group was, in the case of 1, 2, 3, 4, 5, equal to 10. For 7, 9, 11, 13, 15, it was 40. Now if we were to throw both sets of numbers together to form a new set of 10 numbers we would have 1, 2, 3, 4, 5, 7, 9, 11, 13, 15.

The average of these 10 numbers is 7 so their variance is one tenth of the sum $(7\text{-}1)^2 + (7\text{-}2)^2 + (7\text{-}3)^2 + (7\text{-}4)^2 + (7\text{-}5)^2 + (7\text{-}7)^2 + (7\text{-}9)^2 + (7\text{-}11)^2 + (7\text{-}13)^2 + (7\text{-}15)^2$, or 21. The sum of squares in this case is 210—much more than the sum of squares for the two sets of five added together, which is ten plus 40, or only 50. The interesting thing is that we can find the missing 160. If we substitute for the first five numbers their average repeated five times (there were five numbers in the set), and similarly replace the last five by *their* average repeated five times, we have: 3, 3, 3, 3, 3, 11, 11, 11, 11, 11. The sum of squares of this set is the missing 160.

This illustrates how the sums of squares of a set of numbers can be divided up. If a set of numbers is separated into two parts, the sum of squares of both together is equal to the sum of squares of the first, plus the sum of squares of the second, plus the sum of squares of the averages of each set weighted by its number of members.

Something else comes out of this example. Three different sums of squares and three different variances are involved.

Now if the first and second sets of numbers were both chosen at random from the same much larger population of numbers, we would expect their variances to be not very different from one another. Both would tend to be like the variance of the population from which they come.

We might also expect the variance of the composite of the two sets to provide an even better estimate of the population variance. Better, because it would be based on twice as many observations. Again, our small sets have averages and these averages will be scattered around the population average. *Their* variance, too, we would expect to be based on the population variance in some way. If it is, then it offers still another way of estimating the population variance. These two ways of estimating the population variance—from a composite of samples and from the averages of those samples—are quite different. Clearly both estimates should be roughly the same, for the population variance is some fixed number. If they are not at all the same, then we must conclude that the hypothesis that the two different samples came from the same population is probably wrong.

Intuition tells us that here, somewhere, is a test which will tell whether different sets of numbers come from the same or different sources. What follows in this article is a more particular and elaborate phrasing of this testing procedure and of some of the ways of using it most efficiently. The procedure is known as the analysis of variance; the arrangement of situations where it can be used is generally called the design of experiments.

### *Samples and Populations*

When we take samples from a large population and take measurements on each member of the sample, we expect these measurements to reflect those we would find in the total population. If we want to investigate the smoking habits of the nation, for instance, we interview only a sample of the population and then consider our sample as somehow reflecting the behavior of the whole population. In such a case we expect, if the sample has been well chosen, that averages calculated from the sample will be good estimates of averages in the population. We do not expect them to be exactly equal to the population average but they are usually close. The larger the sample, the closer we expect them to be. Moreover the amount of variation we find in the sample we would take as a good estimate of the variation in the population.

More precisely we can say that we expect sample averages to estimate population averages and sample variances to estimate population variances. Since the sample averages are scattered around the population average, we can and do speak of the variance of these sample averages as "the sampling variance of the mean." This sampling variance of the mean we expect to be smaller as the number in the sample gets larger. It can easily be shown that we may expect it to be one $n$th of the variance in the population, where $n$ is the number of members in the sample. An obvious exception to this rule occurs when we expand our sample so that it eventually becomes the population itself. At that stage the sample average *is* the population average—there is no sampling variance of the mean. When the sample is large enough to be a significant proportion of the population, say as much as a tenth, we make a simple correction (by multiplying it by the proportion of the population *not* included in the sample), but for present purposes we can suppose that the population is very large compared with the sample.

In computing an estimate of the variance of a population from a sample, we need a formula which will give us numbers that will in the long run (on enough samples) average out to the right value; that is, we need an unbiased estimator. Also we want one that gets closer and closer to the right value as we take larger and larger samples; this we call the property of consistency. Then we

also need some way of telling how accurate our estimate is likely to be say nine times in ten, or 99 times in 100; that is, we want to know how the estimate will be distributed about the true value.

It can be shown that the sample variance is an unbiased, consistent estimate of the population variance. There is no more satisfactory way of estimating it.

So far we have been discussing a statistic—the variance. A statistic is just an index calculated from data, any data. We can calculate the variance, for example, from any set of figures no matter what the source of them. But if we want to make inferences about data we have to have some expectation of how the statistic is going to behave. These expectations we get from probability theory, a completely abstract mathematical discipline, which studies the distribution of *random variables*, and which has nothing to say about statistics. The trick is to find a random variable that we think, for various reasons, should describe the way in which our statistic behaves and whose probability distribution is known. Then we find the chance of getting a value of the random variable which departs as much from its expected value as the statistic we have.

Take, for instance, the ratio of the population variance estimated from the sample to the true population variance. The estimate from the sample should be fairly close to the true population variance so the ratio should be fairly close to one. Sometimes it will be rather more than one, sometimes rather less, but generally it is less likely that the ratio will be far from one than close to it.

Tables of the chi-square distribution, one of the best known in probability theory, list the probabilities of getting, by chance alone, values of this ratio beyond levels as great as or greater than one, and as little or less than one. Suppose we take a sample of ten members from a large population and find the sample variance half again as large as the population variance, we could find in a chi-square table that this happens only as often as one time in four by chance alone. The ratio we find from calculating the population variance estimate divided by the true population variance is the statistic chi-square. The table we look in gives the chance of obtaining various values of the random variable chi-square. The statistic chi-square is calculated by the computational formula:

$$\frac{\text{population variance estimator from sample}}{\text{true population variance}}$$

Obviously variances of large samples are better estimates of the population variance than variances of smaller samples. The ratio of the sample to the population variance will more likely be near one with large than with small samples. In other words the chi-square distribution differs for different sample sizes.

### *Degrees of Freedom*

But this statement needs a slight modification. It is not the actual sample size that is important, but the number of *independent* members in it. If we took a sample of 100 store inventories from a retail audit and before tabulating them made up ten more, all identical with one another, each being the average of the 100 stores already reported, we would have 110 returns to analyze. But there would, in fact, be only 100 independent returns.

Similarly when we make an estimate of the population variance from a sample, if we knew the mean of the population we could calculate the estimate of the variance treating every member of the sample as independent. If, however, we use the sample to estimate the population mean first, we assume that there is an equation connecting the members of the sample with this fixed value, the mean, in the population. This is equivalent to losing one independent member of the sample. Imagine, for instance, that we have four members in a sample. Let them have measurements a, b, c, d. Now the sample average is one fourth of $a + b + c + d$. If we assume that this is equal to the population average (say $\mu$), then $a + b + c + d = 4\mu$.

Now the population mean $\mu$ is assumed known when we calculate the variance. Thus we could do without one member of the sample, say d, because $d = 4\mu - a + b + c$. That is, d could be computed from $\mu$ and the independent values a, b and c. When the sample values are bound together by any equation like this, we say they are subject to a *constraint*. The number of independent values left in the sample is called the number of degrees of freedom and the variance estimator is always the sum of squares divided by this number of degrees of freedom.

### *Three Estimates*

When several samples are drawn from the same population we can estimate the variance of the parent population three ways. First we could put all the separate samples together and calculate their variance. Next we could consider just the means of each sample. If there were k samples, we know the sampling variance of the means is their sum of squares divided by (k-l), the number of degrees of freedom of this sum of squares. Since this is only

one nth of the population variance, where there are n members in each sample, we can get another estimate of the population variance by multiplying this sampling variance of the mean by n. Finally we could get a third estimate of the population variance from the variances of each of the separate samples. These we would pool together by adding the sums of squares from each sample and dividing by the appropriate number of degrees of freedom.

We have mentioned that the sums of squares are additive; so are the degrees of freedom. For example, if we calculate the variance of the numbers 1, 2, 3, we have $\frac{(1\text{-}2)^2 + (2\text{-}2)^2 + (3\text{-}2)^2}{2} = 1$. The two in the denominator is the number of degrees of freedom of the sum of squares in the numerator. Similarly, the variance of 5, 6, 7, is 1. If we put both sets together we have a new set 1, 2, 3, 5, 6, 7 and its variance is

$$\frac{(4\text{-}1)^2+(4\text{-}2)^2+(4\text{-}3)^2+(4\text{-}5)^2+(4\text{-}6)^2+(4\text{-}7)^2}{5} = \frac{28}{5},$$

where the five in the denominator is the number of degrees of freedom of the sum of squares in the numerator.

If, however, we had calculated an estimate of the population variance from a knowledge of the variance of the two separate groups we should have $\frac{(2+2)}{4} = 1$, where the four in the denominator is the number of degrees of freedom of the sum of squares in the numerator. Here two degrees of freedom have been lost; one was used in calculating the average of 1, 2, 3 and a second was lost in calculating the average of 3, 4, 5. To get this pooled variance of the two samples we add the sums of squares from each and divide by the sum of their degrees of freedom.

Then we can make a third estimate of the population variance from the variance of the mean of the two separate groups around the mean of both together: $3\left[\frac{(2\text{-}4)^2 + (6\text{-}4)^2}{1}\right] = 24$. Here again, the one in the denominator is the number of degrees of freedom of the sums of squares in the numerator. This overall variance was calculated to be $\frac{28}{5}$ and we see that both numerator and denominator can both be partitioned like this:

28 (the overall sum of squares) =
2 (the sum of squares in the first sample) +
2 (the sum of squares in the second sample) +
24 (the sum of squares of the means of the samples)

5 (the overall number of degrees of freedom) =
2 (the degrees of freedom from the first sample) +
2 (the degrees of freedom from the second sample) +
1 (the degrees of freedom from the sample of two means)

We noted that the chi-square shows how often various ratios of sample variance to population variance occur, and that the probabilities differed for different numbers of independent members of the sample. More succinctly, we can say that there is a chi-square distribution for each different number of degrees of freedom. Tables of this distribution appear in most statistical textbooks. They show how, when the number of degrees of freedom becomes very large as it does in large samples, the ratio of sample variance to population variance becomes very close to one and there is very little chance of getting a ratio as far away from one as say, .75 or 1.5.

Different samples from the same population usually give slightly different estimates of the population variances. Since the two sample variances are estimating the same thing, we would expect them to be of about equal size and their ratio, therefore, close to 1. One way of looking at it is this: Let $V_1$ and $V_2$ be two different sample variances used to estimate the population variance. The ration of $V_1$ to $V_2$ is $V_1/V_2$. Now suppose that the population variance is V, then we know the probability of getting various values of $V_1/V$ for this ratio follows the chi-square distribution. We also know the probability of getting the various values of $V_2/V$, for this ratio, too, follows a chi-square distribution. Each ratio is associated with a certain number of degrees of freedom depending on the number of independent members in the samples used to calculate $V_1$ and $V_2$. Now the ratio $V_1/V_2$ is equal to $V_1/V$ divided by $V_2/V$ so we see that the ratio of two sample variances drawn from the same population is the ratio of two quantities both of which have a chi-square distribution. Such a ratio is said to have an F distribution, another one of the half dozen best known probability distributions in the theory of probability. There will be a different F distribution for different numbers of degrees of freedom for both $V_1$ and $V_2$. In order to check how likely it is that a particular ratio of $V_1$ to $V_2$ should occur by chance alone, we have to know the degrees of freedom of $V_1$ and $V_2$. We can then look up tables of the F distribution. It provides us with a means of comparing sample variances and is the testing instrument of analysis of variance.

### *Which Soup Tastes Best?*

A few simple examples will show how the technique is used and will indicate how, with a little ingenuity, it can be extended. Suppose six members of a small taste-testing panel are asked to rate four

samples of soup for flavor. They give their ratings on a scale running from zero to ten, like this:

| | SOUPS | | | |
|---|---|---|---|---|
| *Judges* | *W* | *X* | *Y* | *Z* |
| A | 4 | 6 | 5 | 7 |
| B | 5 | 7 | 7 | 4 |
| C | 4 | 5 | 6 | 7 |
| D | 6 | 4 | 2 | 4 |
| E | 6 | 5 | 6 | 5 |
| F | 5 | 4 | 6 | 6 |
| Average | $\frac{30}{6}$ | $\frac{31}{6}$ | $\frac{32}{6}$ | $\frac{33}{6}$ |
| Sum of squares | $\frac{24}{6}$ | $\frac{41}{6}$ | $\frac{92}{6}$ | $\frac{57}{6}$ |

The question is: are all these soups alike to the raters, or do they differ?

One approach to this problem, especially if we were comparing just two soups, would be to calculate the average rating given to each soup and use a standard statistical test to find whether the average ratings differed significantly. The kind of answer such a test gives is that two soups differ in their average ratings by an amount that would only be found as often as one time in twenty by chance alone. The trouble is that if we try this out for all pairs of soups there are six comparisons to make, and things that happen only one time in twenty by chance alone, start happening rather more often.

The analysis of variance approach is different. It consists of seeing the six ratings as a sample of ratings. There are then four such samples, one for each soup, and six members, one for each rating given, to each sample.

The sum of squares for the average ratings of each soup is $\frac{8047}{144}$ and weighted by the number of ratings composing each average is $\frac{8047 \times 6}{144}$. The comparison of the estimates of the population variance provided by the variance for each soup and the variance of the average ratings is then laid out in this form:

| *Source of Variation* | *Sum of Squares* | *Degrees of Freedom* | *Variance Estimate* |
|---|---|---|---|
| Between soups | $\frac{5}{6}$ | 3 | $\frac{5}{18} = .2778$ |
| Within soups | $\frac{214}{6}$ | 20 | $\frac{214}{120} = 1.7833$ |
| Total | $\frac{219}{6}$ | 23 | $\frac{219}{138} = 1.5870$ |

This illustrates the basic method of analysis of variance and, incidentally, the standard method of presenting it. Before considering how this could be complicated to deal with harder questions it is appropriate to explain some assumptions that have been made implicitly.

In the first place, the test just described supposes that the soups each had some kind of true rating. But preventing us from having this true rating given to each soup each time are the perversity of the judges and the vicissitudes of the experimental situation which together add a random element to each true rating. This random element is assumed to afflict all the ratings of every judge and every soup. It will vary in the amount it affects each figure but the degree to which this random element varies, and the frequency which it assumes various levels, is assumed to be described by the normal distribution. The F test assumes this is the case, but it is what statisticians call a "robust" test—one not very misleading even when the data being examined depart somewhat from the assumptions made about them.

Finally, we guess that the four "true" ratings of the soups are all the same and then we see how likely it is that we should get variance estimates which differ as much as they actually do.

Imagine, as the underlying mathematical model assumes, that the rating given to each soup is really a fixed number and that, unfortunately, these fixed numbers are veiled from sight by random additions and subtractions which affect all our data. The actual ratings we have are, according to this model, the "true ratings" seen, as it were, through a glass darkly.

A slightly more complicated mathematical model would take into account the fact that each judge might have his own tendency to be mean or generous with his ratings. The actual observed ratings would then be a composite of three parts: the particular soup being rated, the particular rating and the random variable obscuring these basic influences. This can be illustrated by the kind of data that the mathematical model assumes and the kind that we actually find in practice.

### *Which Copy Is Best?*

Suppose we used four different copy platforms to advertise a product and that each was tried in each of five cities, only one approach being tried in one city at a time. Suppose, too, that we have assumed that there is no pronounced order effect of copy platforms—that the effect of each on its successor is negligible. The mathematical model, or at least, one mathematical model, of analysis of variance assumes

that the "real" sales results, if only we could see them, would be like this:

| Copy Approach | City A (+10) | B (−8) | C (+12) | D (+2) | E (−16) |
|---|---|---|---|---|---|
| I (+3) | +13 | − 5 | +15 | +5 | −13 |
| II (−2) | + 8 | −10 | +10 | +0 | −18 |
| III (−2) | + 8 | −10 | +10 | +0 | −18 |
| IV (+1) | +11 | − 7 | +13 | +3 | −15 |

In this table the numbers in the main body of the table represent units sold, perhaps in thousands, above or below the average for all five cities and four copy platforms. The numbers in parentheses under each city and copy platform indicate the effect of that city or platform. Whatever copy approach is used in city A, for example, will sell ten units more than in the average city, and 20 units more than in city B. Similarly, Platform I will, city for city, always sell five more units than the average platform in that city and five units more than Platform II.

### *The Veil of Reality*

Unfortunately, this is the kind of data that rarely turns up, least of all in marketing and advertising experiments where all sorts of complicating influences obscure these straightforward effects. In practice we might observe the following kind of information, which is no more than the previous table with each observation infected with a normally distributed random variable, and an overall average sale of 50 units in all cities with all campaigns.

| Copy Approach | City A | B | C | D | E |
|---|---|---|---|---|---|
| I | 62 | 46 | 56 | 55 | 37 |
| II | 59 | 41 | 59 | 48 | 33 |
| III | 58 | 40 | 61 | 51 | 32 |
| IV | 63 | 43 | 58 | 54 | 34 |

The data in this table are much less informative. It is impossible to see, simply from inspection, whether the copy approaches differ significantly from one another.

The method of analysis, however, is simple and much as before. This time, instead of just computing the variances within columns and the variances between columns of the table, we add a new dimension. We compute the variance between the cities, that is, between the row means; then we compute the variance of all the figures in the table. Next by subtraction of the sums of squares and degrees of freedom of the rows and of the columns from the total sum of squares and degrees of freedom, we have a "residual" sum of squares and degrees of freedom. From these we can compute an estimate of the population variance which does not take into account the row (platform) or column (city) effects. To determine whether the platforms differ significantly in effect, we compare the estimate of the variance from between rows with the residual variance; to determine whether different cities have a significantly different effect we compare the variance estimate from between cities with the residual variance. The actual analysis would look like this:

| *Source of Variation* | *Sums of Squares* | *Degrees of Freedom* | *Variance Estimate* | *F Ratio* |
|---|---|---|---|---|
| Between cities | 1990 | 4 | 497.50 | $\frac{497.5}{4.93} = 100.9$* |
| Between copy approaches | 35.8 | 3 | 11.93 | $\frac{11.93}{4.93} = 2.42$† |
| Residual | 59.2 | 12 | 4.93 | |
| Total | 2085 | 19 | | |

* The variance between cities is so large that it could have occurred by chance alone less than one time in 1,000. Different cities affect sales significantly.

† The variance between copy approaches could have occurred about one time in ten by chance alone. We would conclude from this evidence that the copy approach did not significantly influence sales.

In this table the "Residual" variance has been obtained by subtracting the sum of the between-city and between-platform variance from the total variance.

The same assumptions have been made in this analysis as in the earlier one. First, that the effects, in this case of city and copy platform, have a simple additive effect. The fact that a particular platform might be just right for one of the towns is the kind of effect that would violate this assumption altogether and cause this particular analysis to be quite misleading. Then it is assumed that the city and platform effects have all been obscured by a normally distributed random variable. If these assumptions are satisfied, the method offers a speedy means of judging whether an effect is important and assessing the degree to which it is.

This second example has shown an extension of the technique to deal with two influences acting simultaneously. The procedure can be extended much further and with ingenuity many experiments can be combined in one large-scale plan.

A simple but well known elaboration of the basic technique is to test the effect of three different influences at once. The type of design most often employed is the Latin Square. The basic principle is that the rows and columns of the table, such as we have employed in the example of copy ap-

proaches and cities, shall each represent a hypothesized influence. But another effect is also considered by applying it selectively to the observations represented by the cells of the square table.

We might, for example, wish to test the effect of four different advertising themes, four different price levels and four cities all at once. As before, we could lay out a square table which would describe not only our observations but our way of getting them. Each column of our table will represent a different city; each row a different advertising theme. The problem then, is to experiment with price. This we do by applying each of the experimental prices in each city and with each advertising theme such that any one price appears once only in any city or with any theme. Then in analyzing the results we find the effect of price by adding together the sales from all cities with the same price level. The following table illustrates the procedure.

| | *City* | | | |
|---|---|---|---|---|
| *Theme* | *1* | *2* | *3* | *4* |
| I | A | B | C | D |
| II | B | A | D | C |
| III | C | D | A | B |
| IV | D | C | B | A |

In this table the price levels are denoted by A, B, C and D. Each appears once only in each row and column. Extra observations provided, for example, by an extra city, can be used, of course, but the table presents the minimum design required to test the third effect. The square in the illustration is not the only way of arranging prices within cities and themes. There are over 500 other arrangements possible.

Nor does the Latin square exhaust ingenuity in this direction. We could add another effect, say that of different packaging, and test that at the same time. If we denote few different packages by the Greek letters $\alpha$, $\beta$, $\gamma$, $\delta$, we can test their effect with the experimental design

| | *City* | | | |
|---|---|---|---|---|
| *Theme* | *1* | *2* | *3* | *4* |
| I | $\alpha$A | $\beta$B | $\gamma$C | $\delta$D |
| II | $\gamma$B | $\delta$A | $\alpha$D | $\beta$C |
| III | $\delta$C | $\gamma$D | $\beta$A | $\alpha$A |
| IV | $\beta$D | $\alpha$C | $\delta$B | $\gamma$A |

Here the method of distributing the different packages and prices is self-evident. The important point is that every possible combination of price and package occurs. We have superimposed two Latin squares, one denoted by Latin letters, the other by Greek letters so that every pair appears once only. What is more, yet another treatment, say product variation, could be added, so that every pair of variations in product, packaging and price will occur once and only once.

The superimposition of Latin squares in this way is not always possible. We know it can be done if the number of treatments is a prime number or 4, 6, 8 or 9.

This is merely one of several ways of extending the simple basic techniques to serve elaborate ends. Whatever the complexity of the design employed, two things remain the same. First, the method of analysis will be the subdivision of sums of squares among different influences in order to make estimates of the population variance. Second, the analysis assumes that each influence acts independently and additively, and that the chance variations which appear in the experiment are normally distributed.

*The test of significance does no more, and attempts no more, than to answer the straightforward question, "Could these samples have been drawn from the same population?" It calculates a probability. If the probability is very small the answer is "No." If it is not so small as to reach the level of significance required, the answer is "Yes, they could." The answer is never "Yes, they must have been."*

—R. A. FISHER

# RESEARCH IN REVIEW

GWYN COLLINS, THORNTON C. LOCKWOOD and GEORGE M. SHIREY, JR.
ARF Staff

## Longe-Range Effects of TV Viewing: II

**Wilbur Schramm, Jack Lyle and Edwin B. Parker. *Television in the Lives of Our Children.* Stanford: Stanford University Press, 1961, 324 pp., $6.00.**

In advertising research we sometimes forget that most advertising is carried by a medium whose function is to entertain, inform and educate, as well as to advertise. In this country and elsewhere, many people have questioned whether media in general, and television in particular, are adequately carrying out these functions.

Critics have been especially concerned about the effects of television on children, since children are more available to the small screen than to print media and are obviously attracted by it. Until 1958 the critics went unanswered, and of course, unsubstantiated as well. Then Himmelweit, Oppenheim and Vince published *Television and the Child,*[1] the results of their studies on the effect of television on British children. Schramm, Lyle and Parker's book is a similar report but deals with television and children in the United States and Canada.

The wide interest in this topic is well served by the authors who have presented the results in two conceptual parts. The first has a minimum of tables and charts, and the statistics are so well woven into the text that they pass almost unnoticed; the second part gives a detailed description of the research methods, the statistical findings, and an extensive and thoroughly annotated bibliography. The appeal of this book is such that it has been reviewed in the popular daily press in Britain.

It is based on 11 studies made between 1958 and 1960—two in San Francisco, five in the Rocky Mountains, two in Canada, one in Denver and another in an American suburb. Some 6,000 children from the first grade through high school and 2,300 parents, teachers and school officials participated in the studies. The Canadian studies compared two similar communities, one with television, the other without.

As in the earlier British study, Schramm, Lyle and Parker found that television has little effect on the "normal" child. If introduced to television during preschool years, he is usually one grade ahead in vocabulary upon entering school but by the sixth grade has lost this advantage. Television does change his routine of life to some extent, mainly in the emphasis of his leisure time and his degree of contact with other mass media. For example, play time, attendance at movies, reading of comic books and pulp magazines is reduced. As the child grows older, his viewing hours increase until he reaches the sixth or seventh grade, then they fall off slowly throughout high school. The average length of time viewing seems to be about two to three hours per day—not the four or five hours so frequently quoted. Physical and emotional effects of television apparently are slight, but so is the effect of the medium in supplying information. Television, it

[1] Reviewed in the June 1960 issue of this *Journal.*

seems, is not very successful in transmitting new ideas and new knowledge to the child.

One particularly interesting section of this book explores the child's use of television. There are some obvious reasons: entertainment and escape, information, topics of conversation. Less obvious needs are experiences which the authors divide into fantasy and reality. Their use of these terms is analogous to Freud's pleasure and reality principles. Very roughly, the former may be considered as experiences providing emotional relief, the latter being experiences offering positive orientation to the real world. Television, it seems, is meeting the needs for fantasy but not providing the necessary reality experiences.

In monitoring four commercial stations in one metropolitan area during the "children's hour"—4:00 to 9:00 p.m.—the authors found 18 per cent of programs made up of animated cartoons, five per cent ancient slapstick movies, 10 per cent feature movies, 13 per cent westerns, 11 per cent crime or whodunits, and two per cent travel. During the same time period only 15 of 100 hours were devoted to reality programs, and half of these were news reports. About 22 per cent of the time a child could find a realistic program on at least one of the four commercial stations. In contrast to these, a noncommercial station offered imaginative programs for the very young, and reality programs for the more intelligent teenagers and adults. The effects of television will, of course, vary from child to child. The authors consider its impact on children of differing mental abilities, social norms, social relations, and so on.

In summing up, the authors pose some interesting questions for various groups in the population. Addressing the broadcasters, Schramm, Lyle and Parker, who stress they have been careful not to push the panic button, ask if such large doses of violence and excitement are necessary; they ask whether more challenge could be offered to bright children; and whether the picture of adult life now offered could be made more appropriate to their needs. Parents are asked to consider whether they are doing enough to make their children feel secure, loved, "belonging"; whether they are guiding their children to select the more stimulating experiences on television; and whether they are doing all that is possible to improve what is now being offered. Schools are asked to direct the child to the reality experiences of television and make use of television *in* the school, and the Government is urged to encourage the stations to provide better programs. Researchers are encouraged to undertake intensive and long term studies of television in the lives of children.

This is a book to be commended to all these diverse groups—and to all who are interested in the future of television.

—GEORGE M. SHIREY, JR.

# Measurement Economy At Any Cost

**James E. Brinton. Deriving an Attitude Scale from Semantic Differential Data. *Public Opinion Quarterly*, Vol. 25, No. 2, Summer 1961, pp. 289-295.**

**Jack Lyle. Semantic Differential Scales for Newspaper Research. *Journalism Quarterly*, Vol. 37, No. 4, Autumn 1960, pp. 559-646.**

Both authors have used the semantic differential to measure attitudes, Brinton those of college students toward capital punishment, and Lyle those of subscribers toward newspapers. Together the articles suggest an important new approach to measuring advertising effectiveness.

Brinton selected 16 polar adjective scales and administered them to three groups of college students. A seventeenth scale obtained a rating of the respondent's over-all feeling about capital punishment, from strongly in favor to strongly against. The five adjective pairs showing the greatest mean difference between the pro and anti groups (as distinguished by the seventeenth scale) were then arranged as a Guttman scale and satisfied the criterion for high reproducibility, indicating unidimensionality.

From an elaborate adjective-hunting expedition, Lyle selected 32 pairs, adding 25 from Osgood's original semantic differential scale. The resulting 57-item battery was administered to two groups of undergraduate students, who also answered 14 items from the Stanford Test, an earlier measure of attitudes toward newspapers, tapping six unidimensional Guttman scale areas. Scores on the two tests were highly correlated.

Lyle selected 20 of the 57 items which represented a wide range of response. He hypothesized that these 20 items covered five basic dimensions of attitudes to newspapers. The first three—evaluation, potency and activity—were suggested by Osgood, a pioneer in semantic differential measures. These

three, he thought, were common factors underlying a wide range of attitudes. In his study Lyle added two more dimensions which he saw as particularly appropriate to newspapers: bias and accuracy.

After administering these 20 items to subscribers to the *Palo Alto Times,* a factor analysis yielded five factors. Three of these he found reasonably easy to label: newsworthiness, bias and accuracy.

Brinton's paper is a methodological one which explains how to get a Guttman scale from a semantic differential. The object according to Brinton is unidimensionality. Lyle's purpose appears to be to develop a short profile measuring device.

Thus both authors seek economy. Yet Brinton's five-item scale is validated by a one-item scale. Why five items instead of one? Lyle starts with 57 items, reduces them to 20 and validates them with someone else's 14-item scale. What kind of economy is this?

Many papers, nowadays, breed new attitude scales from old. They correlate, mutilate and twist one scale into another. But this is all paper shuffling. The real question is whether these scales relate to any kind of measurable behavior at all.

Attitude scales are useful to the advertiser if they help him predict the probable outcomes of various advertising alternatives. But what interpretation should be given to a plus 5 on a "good-bad" scale? It seems extremely doubtful that any advertiser can be persuaded to study scale scores when he knows so little about their impact on sales, profits or return on investment.

Before the semantic differential can be used with any confidence, considerable development is needed. First, respondents might rate products or brands using a semantic differential inventory. Factor analysis could identify those items representing underlying factors. Unidimensional scaling might enable one score to represent several. Following any such simplifications, differences between brands, and between users and non-users of the same brand, may suggest promising copy approaches. Test ads could be subjected to semantic differential verification prior to publication. With practice it should be possible to work out applications for specific media audiences.

Systematic research is needed. Fortunately, it need not be expensive. Skilled academic researchers often can provide advertisers with exploratory conclusions. These could be turned over to commercial practitioners for development.

—THORNTON C. LOCKWOOD

# Scholarly Beachcombing

**Parker M. Holmes. *Marketing Research Principles and Readings.* Cincinnati, Ohio: Southwestern Publishing Co., 1960. 641 pp. $7.50.**

The preface to this book claims it offers a new approach to the study of marketing research. The book, however, is no more than a selection of 85 readings; a kind of rag bag culled mainly from the popular journalism of advertising, marketing and opinion research. The general character of these articles may be inferred from the fact that three quarters of them come from four magazines: *Industrial Marketing,* the *Journal of Marketing, Sales Management* and *Printer's Ink.* Every article was, of course, appropriate to the journal in which it was found, and some of them would be really good anywhere. Together they are offered as a textbook for the student.

The trauma induced by these bits and pieces forces upon us the question of how market research should be taught, if it can be taught at all. Your reviewer, who has never taught such a course, holds his own views about this with all the dogmatism of ignorance.

Teaching market research is a little like teaching bridge-building. No one starts out just building bridges. There is the long preparation in mathematics, mechanics, strength of materials and so forth before the would-be bridge builder gets 'round to his first bridge. So it is with market research, which after all describes an area of inquiry rather than an academic discipline. Before researching any market the researcher slowly accumulates some knowledge of sampling, of attitude research, statistical analysis and more besides. A collection of capsuled copy testing and nutshelled sampling is no substitute. There is no Instant Market Research to be quickly reconstituted with an instructors manual and a score or so of classroom discussions.

This book will guide the student less to technical knowledge than to an appreciation of what scissors and paste can accomplish. It may encourage him to snip and save for many careful years, until one day his file drawer bursts open and gives the world another book like this one.

It is a sobering thought that if there were only 90 original research papers in the world, from them could be made about 44 million different books of 85 each.

—GWYN COLLINS

# A Compendium Revised

**Claire Selltiz, Marie Jahoda, Morton Deutsch, and Stuart W. Cook. *Research Methods in Social Relations* (Revised in one volume). New York: Henry Holt and Company, 1960, 622 pp., $5.50.**

Three of the four authors of this volume produced an original two-volume edition in 1951. Their purpose was to bring together on an introductory level the considerations which enter into every step of the research process.

Any work attempting to fulfill these objectives faces formidable problems of selection, condensation and organization. The solution chosen for this volume is topical. The first two chapters on the nature of research and research objectives are followed by two others outlining the characteristics of exploratory studies, surveys and experiments. A fifth chapter is devoted largely to validity and reliability. The next five concern data collection techniques. Four concluding chapters consider the organization and analysis of research findings, research reporting form, implementation of research findings, and the interrelationship of theory and research practice. Three useful appendixes discuss the planning and supervising of research, sampling, and questionnaire and interviewing procedures. As each topic is presented the 350-item bibliography is drawn upon to suggest one or more sources for further investigation.

While the revision provides a wider range of illustrative materials than the original, which depended heavily upon research on prejudice, little consideration is given to problems peculiar to specific areas of research. Statistical analysis is not discussed at all.

The scope of some of the topics listed in the table of contents is not as readily apparent as might be wished due to some rather uncommon usage. "Interviews", for example, refers only to personal interviews. "Questionnaires" denotes a mail or group administered interview rather than a schedule or form. The "sociometric method" appears to be nothing more than the "what magazines do you read regularly and like best" approach applied to people. Such terminology will tend to lead those in search of specifics to the detailed index.

While joint authorship has been retained in the revision, chapters attributed to specific authors in the original edition have been merged into the text. The effect is to reduce the over-all level of sophistication to meet the needs of the beginning student. This has led, however, to some over-lap in subject matter. Picture preference techniques, frequently thought of as projective, are found under "Visual Aids in Interviewing" rather than under "Projective Methods". Scaling devices are discussed twice, in Chapters Five and Ten.

The summary at the end of each chapter is usually too brief to summarize the wide variety of techniques discussed therein, with the result that the real summaries tend to get buried in the text without identification. Fortunately, the readability of the volume encourages the search for them. Opinions are succinct and viewpoints are seldom in doubt. There is no other general text in the research field quite like this one.

The book is attractively bound and jacketed. Its over 600 pages are printed on light paper, adding weight but not bulk. The initial impact is one of a Webster or Roget—certainly not an inappropriate objective. The italicized typography, suffering from too little contrast with the text, will prove a minor irritant to those who like to be sure when they've passed an italicized word or phrase.

This work should be extremely valuable as an introductory text, not because it lays out a check list of rules to be committed to memory, but because it invites critical analysis. The meaning and utility of definitions, causality, validity, scaling and experimental design should evoke lively discussion —and not only among introductory students.

—THORNTON C. LOCKWOOD

# All About Those Non-respondents

**Christopher Scott. Research on Mail Surveys, *Journal of the Royal Statistical Society, Series A*, Vol. 124, Part 2, 1961, pp. 143-205.**

This important paper offers some new findings, and a cursory discussion of some old ones, on various aspects of mail surveys. The author, a member of the British Government Social Survey, bases his own findings on five large-scale mail surveys for which he was responsible. All five used probability sampling; three were directed to the general adult population and two to special groups.

The author defines non-respondents unambiguously, as persons to whom an inquiry has been addressed but who have not supplied a usable return. He includes among them the illiterate, the non-existent, the dead and the non-traceable. His response rate, however, is slightly over 90 per cent.

The author reports his hounding of the missing 10 per cent with personal interviewing and his success with two-thirds of them. Unlike other investigators, he is hesitant to compare the results of these personal interviews with his mail returns. He considers it possible that the three per cent who respond to neither mail nor personal approach might have significant characteristics of their own.

By making such fairly safe assumptions as, for instance, that the relation between interview and personal non-response is the same for people under 50 as for people over 50, he is able to say something about characteristics more commonly found among non-respondents. He finds married women and adults who are poor letter-writers or have large families tend to be prominent among non-respondents. He also notes that questionnaires addressed to wives or to adults over 50 years of age tend to be completed by someone other than the addressee.

Mr. Scott next considers response over time. He quotes a formula which predicts that within three days of receiving the first return, half the total response can be expected, within five days three quarters of it and within ten days 90 per cent of it. His own results do not confirm this but apparently for good reason.

More interesting is the *nature* of the response over time. Tabulations are presented of returns received in the first three days, the first week, the first fortnight and the entire period. For most of these characteristics there was an independent population estimate. Two thirds of the characteristics are more accurately estimated from the trend of responses than from the total response. Mr. Scott's findings also suggest that other influences at work were the degree of interest the questionnaire held for the respondent, and the time he needed to complete it. When interest is high the questionnaire is returned quickly; when interest is low it may be delayed.

The author finds that follow-ups significantly increase the response rate. He quotes from the literature the finding that the proportion of returns to each follow-up are about the same as to the initial mailing.

Another question considered by Mr. Scott was the stamped versus metered envelopes. His own investigation and the literature suggests the stamped envelope gives a better response. Even when the stamped return envelope is addressed and the metered envelope is not, the advantage seems to be with the stamped.

Other findings relate to the nature of the accompanying letter, the format, layout and method of reproduction, the color of the questionnaire and the day and date of dispatch; these are difficult to generalize to other occasions. Experiments offering the respondent anonymity were limited, as was the anonymity offered the respondent; they do not suggest that anonymity increases return rate.

The final section of the paper considers the reliability and validity of mail survey responses. Here Mr. Scott makes some sharp comments on the existing literature. He distinguishes between gross error (the total number of individual errors) and the net error (the error in a frequency distribution of all respondents together). Since the several items of gross error often tend to cancel one another out, there has been a tendency to concentrate attention on net error. Mr. Scott also criticizes those researchers who compare two responses from each individual involved in a survey and test as the null hypothesis that the two sets of responses are independent.

Apparently discrepancies between mail and interview responses were slight. Sex, marital status and age were not significantly different. Some discrepancies appear on occupational class but the errors appear to have come equally from interview and mail returns. On attitudinal questions respondents are more pleasant to interviewers than to their mail returns.

This paper is supplemented by an extensive bibliography and a technical discussion of the paper by members of the Royal Statistical Society to whom it was originally read. Altogether it is probably the most thorough and best documented study of the mail survey technique yet written. Near its beginning the author generously but accurately admits that his method of exposition more closely resembles a random walk than a linear progression. It will be unfortunate if this should deter anyone from reading this massively informative article.

—GWYN COLLINS

## Are Campaigns Like Epidemics?

**B. Benjamin, W. P. Jolly and J. Maitland. Operational Research and Advertising: Theories of Response. *Operational Research Quarterly*, Vol. 11, No. 4, December 1960, pp. 205-218.**

Is the explanation of advertising response analogous to that of an epidemic of measles or poliomyelitis?

No irreverence is intended by the analogy. It is suggested as a theory to account for the interval between advertising reception and response (incubation), the interval between responses (immunity and susceptibility), and the variable performance of media (the carrier). This formulation is held tenable because of certain coincidences and similarities in data between the disciplines. The authors, hoping to forestall an obvious criticism, apologize for one such coincidence:

> It must be emphasized that this arrangement of the data, although a matter of subjective judgment, preceded the statistical analysis and was not made to fit that analysis.

Two of these authors, in an article published in 1958 along the same lines, pointed out they were not suggesting that conditions normally present in the analogy really exist, or are even approximated in advertising practice. They might well have restated this qualification in their 1960 article.

In the final three paragraphs of the article the cost utility of predicting advertising responses is briefly touched upon. Here the authors suggest that physical size of space may not be as influential in determining response as its price structure would suggest.

Another aspect of advertising cost is also important. Coupon response is a measurable criterion of advertising performance. In direct advertising the rate of coupon return has often been used to predict total return. Predicting such responses from early returns may play an important role in determining the continued use of specific media and the frequency of such use.

Unfortunately the range of data presented in this article is restricted to coupon responses to a very few print advertisements.

—THORNTON C. LOCKWOOD

# What Viewing Does to Reading

**William A. Belson. The Effects of Television on the Reading and the Buying of Newspapers and Magazines. *Public Opinion Quarterly*, Vol. 25, No. 3, Fall 1961, pp. 366-381.**

As in his prior published work on the subject, the author is concerned with whether television causes other behavior. In this instance he sets out to examine the view that television is responsible for the declining circulation of British national daily and Sunday newspapers, weeklies and monthlies.

The author's test of causality is the degree to which owning or viewing television is associated with other reported activities; the more frequent the association the more evidence of causality. Most readers of the article will readily accept Belson's preliminary observation that the available evidence is not so cut and dried: while the circulation of many newspapers is indeed falling, the circulation of others is rising. Most readers will also question his definition of causality.

Having noted the ambiguities that enter into relating television penetration to over-all circulation trends, Belson still decides to test the effect of television viewing against the buying and reading of individual newspapers and magazines. A study was made of 1,157 adults in Greater London, 288 viewers and 869 nonviewers. Reported newspaper reading and buying were tabulated by viewers and nonviewers and compared. Belson reports that television has caused both increases and decreases in reading and buying of individual publications. One all too typical observation:

> . . . the Daily Mirror gains such considerable support from television that in its absence that paper would have a declining circulation.

He concludes that there must be factors other than just television causing the over-all decline in circulation among newspapers.

Regardless of whether television viewing causes certain reported behavior—or serves simply to weed out certain kinds of already existing behavior patterns—Belson has provided much interesting data. Viewers' and nonviewers' reading and buying are reported for nine morning, ten Sunday and three evening newspapers. He also discusses several other findings:

1. Wherever there is more reading and buying among viewers rather than nonviewers, it tends to be associated with the popular press and its lighter, shorter material. The author speculates that the lower incidence of quality paper buying and reading in viewing homes stems from viewing circumstances that discourage continuous reading.

2. Viewing tends to be associated with increases in reading about TV programs, with increases in the total range and variety of the newspaper items read, and with greater awareness of newspaper advertising.

3. Viewing tends to be associated with increases in reading and buying of daily and Sunday newspapers at the expense of weekly and monthly magazines.

—THORNTON C. LOCKWOOD

## BRIEFLY NOTED

**Harold Bierman, Jr., Laurence E. Fouraker and Robert K. Jaedicke. *Quantitative Analysis for Business Decisions.* Homewood, Illinois: Richard D. Irwin, Inc., 1961, 358 pp., $10.60.**

A quick glance at some of the black arts of operations research—inventory models, game theory, linear programming and all that. Little mathematics except in the appendices.

**Eleanor Blum. *Sources for Reference in the Mass Media, An Annotated, Selective Bibliography.* Urbana, Illinois: Institute of Communications Research, University of Illinois, 1961, 46 pages. Free.**

A bibliographic bargain. Contains 185 beautifully annotated secondary and primary sources (some unique), grouped in six categories: books and publishing, broadcasting, motion pictures, newspapers and periodicals, advertising and general. Has no index, but all items are "cross-annotated" so as to lead the reader to related items from wherever he may begin. In short supply; an expanded edition is due in spring 1962.

**S. Watson Dunn. *Advertising: Its Role in Modern Marketing.* New York: Holt, Rinehart and Winston, 1961, 621 pp., $8.95.**

A good, readable text. Contains the kinds of information—on the growth of TV, approximate costs of ad space, etc.—which every new recruit to advertising needs for orientation. This book deserves and will doubtless get attention in wider circles than business students and the industry provide.

**John M. Fenton. *In Your Opinion.* Boston: Little, Brown & Company, 1961, 220 pp., $3.95.**

A review of post-war U. S. social history as reflected by the Gallup Poll. Journalistic in style, it makes good light reading.

**Andrew E. Kimball. Increasing the Rate of Returns in Mail Surveys. *Journal of Marketing*, Vol. 25, No. 6, October 1961, pp. 63-64.**

Reports a simple experiment using air mail stamps and air mail permits, personal address and the impersonal "Dear Sir", money incentives and no incentives. By the criterion of return rate, stamps seem better than permits, incentives better than no incentives and the style of address irrelevant.

**John C. Maloney. Portfolio Tests—Are They Here To Stay? *Journal of Marketing*, Vol. 25, No. 5, July 1961, pp. 32-37.**

The author believes measures from portfolio copy tests reflect more ephemeral product interest than power of ads. His debatable conclusion: portfolio tests must go.

**C. H. Sandage (Editor). *The Promise of Advertising.* Homewood, Illinois: Richard D. Irwin, Inc., 1961, 207 pp., $5.50.**

Fourteen statements by as many authorities on various aspects of the role of advertising in society. Idealistic, inspiring, national-purpose-type entries selected with obvious care.

**Morris John Slonim. *Sampling in a Nutshell.* New York: Simon and Schuster, 1960, 144 pp., $3.50.**

A light-hearted survey of sampling. Readable, elementary and accurate within its self-imposed limits, it will give anyone an appreciation of what sampling statisticians do. Probably the funniest book in this field since Darrell Huff's *How To Lie with Statistics.*

**J. Walter Thompson Company. *Population and Its Distribution. The United States Markets* (Eighth Edition). New York: McGraw-Hill, 1961, 471 pp., $24.00.**

This well known handbook of U. S. marketing statistics is now based on the 1960 Census returns. Speedily produced, it is an enlarged, improved version of earlier volumes in the series.

**W. T. Tucker and John J. Painter. Personality and Product Use. *Journal of Applied Psychology*, Vol. 45, No. 5, October 1961, pp. 325-329.**

A short, well-written report of the relationships between four personality scores and reported frequency of product use. Subjects were 101 college students. Of 36 correlation ratios, 13 were significant at the .05 level or better. A forthcoming ARF methods research report will describe similar relationships for a panel of some 3000 U.S. households.

**Martin Weinberger. Does the "Sleeper Effect" Apply to Advertising? *Journal of Marketing*, Vol. 25, No. 6, October 1961, pp. 65-67.**

An undigested, uncritical account of research on one aspect of attitude change resulting from deliberate persuasion.

**Robert V. Zacher. *Advertising Techniques in Management.* Homewood, Illinois: Richard D. Irwin, Inc., 1961, 618 pp., $10.**

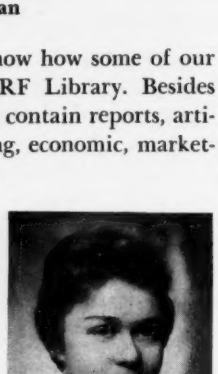
Elementary treatment of advertising organized around familiar topics. Emphasizes producing advertisements and advertising management. Profusely illustrated.

# PUBLICATIONS RECEIVED

**NAOMI BORETZ**

**ARF Librarian**

You may be interested to know how some of our members make use of the ARF Library. Besides books and periodicals, our files contain reports, articles and clippings in advertising, economic, marketing and social science research.



Recently a member spent several hours in the reading room compiling a bibliography on the introduction of new products. It included items on product planning and development, market tests, product testing, test cities, point-of-purchase advertising and product surveys. He later wrote that in two days spent in the ARF Library, he had accomplished more than he had in two weeks elsewhere.

An agency member visited just to get acquainted with the Library, was impressed by its size, and found three books he wished to borrow immediately.

Six articles and a list of suggested readings on the subject of brand loyalty were sent in one package to a member. His research department needed some information immediately: the items were collected and a short bibliography compiled in about an hour. The firm sent a messenger for the materials and it was in their hands the same afternoon.

Members outside the New York City area have often called or written about material they wanted to see. We may not mail any of our books but only clippings and reports. A media member called recently for information on outdoor advertising and was loaned 12 relevant articles by return mail.

The Library keeps an up-to-date file on research organizations, their services, and their publications. Many visitors have referred to this file and found it helpful.

The reference shelves hold a complete set of ARF publications, including those out-of-print, such as the 1937 study on window-display, and the *Continuing Study of Newspaper Readership* (1939-1951).

The Library is open from nine to five Monday through Friday. We advise members who wish to visit to call or write beforehand to make sure that the material they need is available.

We have written to 31 editors offering to exchange our *Journal* for theirs. This offer was accepted by the following:

*Behavioral Science*. Mental Health Research Institute, University of Michigan. Quarterly.

*Incorporated Statistician* (British). Association of Incorporated Statisticians, Ltd. Quarterly. Editor: W. H. B. Carey.

*International Abstracts in Operations Research*. Operations Research Society of America. Editor: Herbert P. Galliher.

*Journal of the American Statistical Association*. American Statistical Association. Quarterly. Editor: Clifford Hildreth.

*Journal of Applied Psychology*. American Psychological Association. Bi-monthly. Editor: Kenneth E. Clark.

*Journal of Business*. University of Chicago. Quarterly. Editor: Irving Schweiger.

*Journal of Educational Research*. Dembar Publications, Madison, Wisconsin. Nine times yearly. Editor: A. S. Barr.

*Journal of Marketing*. American Marketing Association. Quarterly. Editor: Steuart Henderson Britt.

*Journal of Social Issues*. Institute for Social Research, University of Michigan. Quarterly. Editor: Dr. Robert Chin.

*Merrill-Palmer Quarterly of Behavior and Development*. Detroit, Michigan. Quarterly. Editor: Dr. Marvin L. Hoffman.

*Perceptual and Motor Skills*. Southern Universities Press, Missoula, Montana. Bi-monthly. Editors: R. B. Ammons and C. H. Ammons.

*Psychologia* (Japanese). Kyoto University. Bi-monthly. Editor: Koji Sato.

*Psychological Record*. Denison University. Quarterly. Editor: Irvin S. Wolf.

*Public Opinion Quarterly*. Princeton University Press. Quarterly. Editor: Frederick F. Stephan.

*Quarterly Review of Economics and Business*. Bureau of Economic and Business Research, University of Illinois. Quarterly. Editor: Marvin Frankel.

*Sociometry*. American Sociological Society. Quarterly. Editor: Leonard S. Cottrell, Jr.

Since the appearance of the September issue of the *Journal* the Library has received a number of suggestions concerning bibliographies. Other members are encouraged to request additional topics for literature searches. When we have a sufficient number of suggestions we will poll our membership to ascertain which would be of most interest for an extensive, annotated bibliography.

Publications acquired since the last issue include:

American Management Association. *Market Research in International Operations*. A.M.A. Management Report No. 53. New York: American Management Association, 1960. 88 pp. $2.25.

American Management Association. *Operations Research Reconsidered*. A.M.A. Management Report No. 10. New York: American Management Association, 1958. 143 pp. $3.75.

Association of National Advertisers. *Magazine Circulation and Rate Trends, 1940-1959*. New York: Association of National Advertisers, 1960. 188 pp. $15.00.

Baker, Stephen. *Visual Persuasion: The Effect of Pictures on the Subconscious.* New York: McGraw-Hill Book Company, 1961. $13.50.

Berlo, David K. *The Process of Communication: An Introduction to Theory and Practice.* New York: Holt, Rinehart and Winston, 1960. 318 pp. $4.25.

Bierman, Harold, Jr. Lawrence E. Fouraker, and Robert K. Jaedicke. *Quantitative Analysis for Business Decisions.* Homewood, Illinois: Richard D. Irwin, 1961. 358 pp. $10.00.

Blum, Eleanor. *Sources for Reference in the Mass Media. A Bibliography.* Urbana, Illinois: Institute of Communications Research of the University of Illinois, 1961. 46 pp. Free.

Briggs, Asa. *The Birth of Broadcasting. The History of Broadcasting in the United Kingdom.* Volume I. London: Oxford University Press, 1961. 425 pp. $10.00.

British Market Research Bureau, Ltd. *Western Europe: Facts on a Market of 350,000,000 People.* Fifth edition. London: British Market Research Bureau, 1961. 215 pp. Free.

Carpenter, Robert N. *Guidelist for Marketing Research and Economic Forecasting.* A.M.A. Research Study No. 50. New York: American Management Association, 1961. 94 pp. A.M.A. members, $2.50, non-members, $3.75.

Clarke, George T. *Bibliography of Advertising and Marketing Theses for the Doctorate in United States Colleges and Universities,* 1944 to 1959. New York: Advertising Educational Foundation, 1961. 28 pp. $2.00.

Colley, Russell H. *Defining Advertising Goals for Measured Advertising Results.* New York: Association of National Advertisers, 1961. 114 pp. $7.50.

Converse, Paul D. *The Beginning of Marketing Thought in the United States.* Studies in Marketing No. 3. Austin, Texas: Bureau of Business Research of the University of Texas, 1959. 89 pp. $1.50.

Crawford, John W. *Advertising Communications for Management.* Boston: Allyn and Bacon, 1960. 38 pp. $7.50.

Dunn, S. Watson. *Advertising: Its Role in Modern Marketing.* New York: Holt, Rinehart and Winston, 1961. 621 pp. $8.95.

Fenton, John M. *In Your Opinion.* Boston: Little, Brown and Co., 1960. 220 pp. $3.95.

Forrester, Jay W. *Industrial Dynamics.* New York: John Wiley and Sons, Inc., and the Massachusetts Institute of Technology Press, 1961. 464 pp. $18.00.

Guest, Lester. *A Longitudinal Study of the Development and Modification of Attitudes.* Research Bulletin No. 15. University Park, Pennsylvania: The Pennsylvania State University, 1961.

Holloway, Robert J. *Marketing Research and Market Planning for the Small Manufacturer.* Minneapolis: University of Minnesota, 1961. 52 pp. $1.50.

Nation's Business. *A Study of the Corporate Personality: How Business Executives Appraise Leading Instrument Manufacturers.* New York: Nation's Business, 1961. 82 pp. Free.

Nielander, William A. *A Selected and Annotated Bibliography of Public Relations.* Fourth edition. Austin, Texas: Bureau of Business Research of the University of Texas, 1961. 58 pp. 50¢.

Reeves, Rosser. *Reality in Advertising.* New York: Alfred A. Knopf, 1961. 154 pp. $3.95.

Sandage, C. H. (Editor). *The Promise of Advertising.* Homewood, Illinois: Richard D. Irwin, 1961. 207 pp. $5.50.

Schramm, Wilbur, Jack Lyle and Edwin B. Parker. *Television in the Lives of Our Children.* Stanford, California: Stanford University Press, 1961. 324 pp. $6.00.

Selltiz, Claire et al. *Research Methods in Social Relations* (revised in one volume). New York: Holt, Rinehart and Winston, 1960. 622 pp. $5.50.

Shoaf, Robert E. *Emotional Factors Underlying Industrial Purchases.* Cleveland, Ohio: Steel, The Metalworking Weekly, 1959. $4.00.

Slonim, Morris James. *Sampling in a Nutshell.* New York: Simon and Schuster, 1960. 145 pp. $3.50.

Taylor, Weldon J. and Roy T. Shaw. *Marketing: An Integrated Analytical Approach.* Cincinnati, Ohio: Southwestern Publishing Co., 1961. 710 pp. $7.50.

Thompson Company, J. Walter. *Population and Its Distribution. The United States Markets.* Eighth edition. New York: McGraw-Hill Book Company, 1961. 471 pp. $24.00.

Thompson, Walter, de Mexico, S. A. *The Mexican Markets.* First edition. Mexico, D.F.: Walter Thompson de Mexico, S.A., 1959. 705 pp. $10.75.

U.S. Bureau of the Census. *Statistical Abstract of the United States, 1961.* Eighty-second edition. Washington, D. C.: Bureau of the Census, 1961. 1,037 pp. $3.50.

*United States Government Organization Manual, 1961-62.* Washington, D. C.: Office of the Federal Register, 1961. 821 pp. $1.50.

Watkins, Julian Lewis. *The 100 Greatest Advertisements: Who Wrote Them and What They Did.* New York: Dover Publications, 1959. 234 pp. $1.25.

Weiss, Lionel. *Statistical Decision Theory.* New York: McGraw-Hill Book Company, 1961. 195 pp. $7.50.

Wilson, H. H. *Pressure Group: The Campaign for Commercial Television in England.* New Brunswick, New Jersey: Rutgers University Press, 1961. 232 pp. $4.00.

Zacher, Robert V. *Advertising Techniques and Management.* Homewood, Illinois: Richard D. Irwin, 1961. 618 pp. $10.00.

EDITORIAL

# The Balance Sheet

Last month a major advertising agency and a major research supplier jointly announced a "scientific breakthrough"—their use of linear programing to select the media schedule to reach that audience which best matches, in certain ways, a desired audience determined by management's marketing objectives.

This announcement revived three tired issues. Will the computer preclude the media buyer's use of creative judgment? Are the data sufficiently accurate to warrant their costly processing into correspondingly accurate (or inaccurate) outputs? Is the value of such processing simply the requirement it imposes on a manager to specify his objectives numerically?

To raise these questions is distressing, not just because they were carefully answered by the agency and its supplier at the outset, but because they have been answered for such a long time. We suspect that their counterparts must have been asked—and adequately answered—on every occasion in the history of science when its findings were applied to decisions of consequence. Is it not likely that the first savage to put wheels on his sledge had to spend hours by the campfire, grunting the answers: no, the cart wouldn't limit his judgment of where to go; no, he didn't really have enough possessions to need a cart to carry them, but he might have some day; yes, you should know where you were going before you started to roll.

The analogy is barbarous in more ways than one, but it does rest on at least three truisms common to both situations.

New hardware will never stifle creativity.

The data we must use are never as good as we would like.

Objectives must be defined before they can be achieved.

Perhaps by dubbing these truisms we can prevent their resurrection as debatable resolutions. The more we debate them, the less time we have to discuss a proposition logically prior to all three: *that we are fast running out of new knowledge from which to draw future technological applications.*

Failure to see the priority of this question stems from the confusion of two different activities. Science adds to the sum of human knowledge; technology applies it to human welfare, be it individual or corporate.

Du Pont president Crawford Greenewalt once made this distinction by calling basic scientific knowledge "a savings account from which we make withdrawals as we convert that knowledge through applied research to new products and processes. As with all savings accounts, bankruptcy lies ahead when withdrawals exceed deposits. To keep our scientific balance in healthy condition, we must insure always that our deposits derived from basic research are never less than our technological withdrawals."

Our distress stems not just from last month's publicized pride in a technological withdrawal, but from the possibility that some may have mistaken it for a deposit.

These deposits are accumulated by *basic* research. The latter is defined by the National Science Foundation as those investigations in which "the primary aim of the investigator is a fuller knowledge or understanding of the subject under study, rather than a practical application thereof."

Where can investigators work who have these aims? Government and university laboratories are the traditional answer, but private industry has also recognized an obligation. In 1958 Robert E. Wilson, former board chairman of the Standard Oil Company of Indiana, said of support for basic research that "industry has rather generally been considered to be the principal villain in the picture, and not without reason. Industry has been using up the stockpile of basic research information faster than it is being added to, particularly since the beginning of World War II."

Today many large companies are working to reverse this trend, but few are adding to the factual stockpile of the disciplines on which the marketing man draws so heavily. Who in marketing research has deposited new knowledge in the technological savings accounts of psychology, statistics, sociology, demography or economics? The best we can adduce are a few methodological innovations, and not all of these have been made public.

A real scientific breakthrough is heralded not by the bang of an advertised technological application, but by the whimper of new-born knowledge. This takes basic research, the only alternative to an abortive intellectual bankruptcy of the embryo science of marketing.

—CHARLES K. RAMOND

# Journal of Advertising Research
## Index to Volume 1: September 1960—December 1961

### ARTICLES


Agostini, J.-M. **How to Estimate Unduplicated Audiences** No. 3, pp. 11-14

Monsieur Agostini has developed a formula which may render unnecessary the laborious tabulation of the unduplicated audiences of combinations of three or more media vehicles. It has been successfully applied to both French and U.S. magazine data.

Appel, Valentine and Milton L. Blum **Ad Recognition and Respondent Set** No. 4, pp. 13-21

Often an ad is evaluated in terms of the proportion of people who claim to have seen it before. The authors find that this claim depends not only on the person's exposure to the ad, but also on the number of magazines he reads and his interest in the product advertised.

Banks, Seymour **Some Correlates of Coffee and Cleanser Brand Shares** No. 4, pp. 22-28

From an analysis of purchases, prices, preferences and promotion of two everyday products, Dr. Banks is able to find some factors likely to be important in explaining a brand's share of the market.

Becknell, James C., Jr. **The Influence of Newspaper Tune-In Advertising on the Size of a TV Show's Audience** No. 3, pp. 23-26

Twenty-one cities were randomly assigned to groups receiving normal, double and no tune-in advertising for a special TV program. Coincidental viewing measures indicated that the ads had no effect on audience size, but may have served to decrease audience variability.

Cahalan, Don **Measuring Newspaper Readership by Telephone: Two Comparisons with Face-To-Face Interviews** No. 2, pp. 1-6

Do telephone and face-to-face interviews obtain roughly the same answers? The author found they did when a particular newspaper readership question was asked of new-car owners in and around New York.

Copland, Brian D. **Exposure and Communication Measures of Outdoor Advertising in Britain** No. 1, pp. 13-17

Two British studies of outdoor advertising have adopted different criteria of effectiveness. The author discusses the pros and cons of each criterion and its application to advertising in other media.

Dalenius, Tore **Treatment of the Non-Response Problem** No. 5, pp. 1-7

"Survey estimates will never be unbiased—too many in the sample can't be reached." Dr. Dalenius deplores such defeatism and suggests three ways of handling non-response.

Done, Arthur A. **Differences Between Advertised and Medical Uses of Drugs** No. 1, pp. 18-21

What a doctor expects from a drug is not always the same as its advertised properties. So finds this study of ten leading drugs as used by a nationwide panel of 1,000 physicians.

Ehrenberg, A. S. C. **How Reliable Is Aided Recall of TV Viewing?** No. 4, pp. 29-31

This British analysis finds claims of viewing television virtually identical for recall periods of from one to seven days.

Fisk, George **A Survey of Case History Advertising** No. 5, pp. 16-21

The case history is being used as copy by more and more advertisers. Professor Fisk reports who they are, why they use it, how much it costs them and what their problems are.

Frankel, Lester R. **How Incentives and Subsamples Affect the Precision of Mail Surveys** No. 1, pp. 1-5

This comparison of two mail surveys has found that a first mailing incentive increased not only the response to that mailing, but also the precision of estimates obtained from three mailings combined. The author goes on to derive a rule for selecting the most efficient proportion of non-respondents to subsample in repeated mailings.

Henderson, Peter L., James F. Hind and Sidney E. Brown **Sales Effects of Two Campaign Themes** No. 6, pp. 2-11

Advertising is only one of many factors influencing purchase. Here is an experiment which took the most significant factors into account and permitted a sales comparison of the themes employed.

Hill, Conrad R. **Another Look at Two Instant Coffee Studies** No. 2, pp. 18-21

Professor Hill suggests that direct and indirect questions both have limitations in consumer research because answers to both are likely to be rationalizations.

Jessen, R. J. **A Switch-Over Experimental Design to Measure Advertising Effect** No. 3, pp. 15-22

The influence of advertising may endure long after its placement. Dr. Jessen describes an experiment which takes account of such delayed effects while measuring sales response to current advertising outlay.

Keswick, Gordon M. and Lawrence G. Corey **A Sensitive Measure of Ad Exposure** No. 6, pp. 12-16

The authors have developed a simple apparatus for measuring ad recognition, and find it discriminates among advertisements to which the reader was exposed much, little or not at all.

Koponen, Arthur **Personality Characteristics of Purchasers** No. 1, pp. 6-12

The needs 8,963 persons expressed in a personality test have been related not only to their respective ages, incomes and locations, but also to their purchases of certain products and brands. Dr. Koponen's analysis of these relationships suggests that they can, if given proper weight, help guide advertising strategy.

Kover, Arthur J. and Seymour Lieberman **Selecting Commercial Spokesmen** No. 5, pp. 22-25

How do viewers rate a TV spokesman they like? In this study, as believable, sincere and down-to-earth—but definitely not cold and dull.

Marder, Eric and Mort David **Recognition of Ad Elements: Recall or Projection?** No. 6, pp. 23-25

In this eight-city study of the parts of an advertisement, the authors find most claimed readership to be a product of the respondent's imagination.

Meissner, Frank **Sales and Advertising of Lettuce** No. 3, pp. 1-10

For several years California growers promoted nationally a new brand of lettuce, using as many as eight media. Ratings of two of these media, along with local retail price and temperature, were closely correlated with per capita sales in 22 cities.

Neu, D. Morgan **Measuring Advertisement Recognition** No. 6, pp. 17-22

The recognition method of measuring ad readership has recently met with heavy criticism. The author draws on his extensive experience with the technique to rebut its critics.

Stevens, Bill and Joel Axelrod **Three Ways to Improve Ad Pre-Tests** No. 6, pp. 33-36

Drs. Stevens and Axelrod advocate three simple principles which, they say, would remove much of the present confusion about the purposes and methods of copy testing.

Stock, J. Stevens **A Comparison of Eight Audience Estimates** No. 5, pp. 9-15

Magazine audiences are usually estimated from a sample's answers to questions. The writer finds few differences among estimates obtained by eight different questions.

Stoetzel, Jean **A Factor Analysis of the Liquor Preferences of French Consumers** No. 2, pp. 7-11

If complex behavior is caused by a few simple motives, they can be found in the patterns of that behavior. Professor Stoetzel illustrates this premise with a factor analysis which tentatively explains consumer preference for nine liquors in terms of their sweetness, price and regional popularity.

Vickers, Ronald **A Pre-Test of Four Australian Paint Ads** No. 6, pp. 26-32

In this detailed case history Mr. Vickers explains how his agency considered both qualitative and quantitative factors in selecting an advertising theme.

Weilbacher, W. M. **The Qualitative Values of Advertising Media** No. 2, pp. 12-17

A medium's qualitative value has been used to mean its audience composition, its editorial personality, or its contribution to advertising effectiveness. The author evaluates these definitions in terms of available research techniques, meaningful comparisons within and between media, and media selection.

Wells, William D. **The Influence of Yeasaying Response Style** No. 4, pp. 1-12

Some people tend to say yes more easily than others, whatever question they are asked. Dr. Wells describes a test which identifies these yeasayers, and reports how they differ from naysayers in various other kinds of behavior.

## LETTERS


Haskins, Jack B. **Comment on [Frankel's] "How Incentives and Subsamples Affect the Precision of Mail Surveys"** No. 2, p. 22

Frankel, Lester R. **Rejoinder** No. 2, p. 23

Maloney, John C. **Comment on Meissner's "Sales and Advertising of Lettuce"** No. 4, p. 32

Meissner, Frank **Rejoinder** No. 4, p. 33


## REVIEWS

Abelson, Herbert I. *Persuasion: How Opinions and Attitudes Are Changed*. New York: Springer Publishing Co., Inc., 1959, 118 pp., $3.75. No. 2, p. 27.

Ackoff, Russell L. Unsuccessful Case Studies and Why. *Operations Research*, Vol. 8, No. 2, March-April 1960, pp. 259-263. No. 1, p. 27.

Adler, Lee. Phasing Research Into the Marketing Plan. *Harvard Business Review*, Vol. 38, No. 3, May/June 1960, pp. 113-122. No. 1, p. 28.

Alsop, Joseph. The Wayward Press: Dissection of a Poll. *The New Yorker*, September 24, 1960, pp. 170-184. No. 2, p. 29.

Bass, Frank M. et al. (Editors). *Mathematical Models and Methods in Marketing*. Homewood, Illinois: Richard D. Irwin, Inc., 1961. 537 pages, $11.35. No. 5, p. 33.

Belson, W. A. Effects of Television on the Interests and Initiative of Adult Viewers in Greater London. *The British Journal of Psychology*, Vol. 50, Part 2, May 1959, pp. 145-158. No. 4, p. 41.

Belson, William A. The Effects of Television on the Reading and the Buying of Newspapers and Magazines. *Public Opinion Quarterly*, Vol. 25, No. 3, Fall 1961, pp. 366-381. No. 6, p. 52.

Belson, W. A. Television and Family Life. *Advancement of Science*, Vol. 16, No. 64, 1960, pp. 349-353. No. 4, p. 41.

Belson, W. A. and C. R. Bell. *A Bibliography of Papers Bearing on the Adequacy of Techniques Used in Survey Research*. London: The Market Research Society, 1960. 52 pp. No. 5, p. 38.

Benjamin, B., W. P. Jolly and J. Maitland. Operational Research and Advertising: Theories of Response. *Operational Research Quarterly*, Vol. 11, No. 4, December 1960, pp. 205-218. No. 6, p. 51.

Bierman, Harold, Jr., Laurence E. Fouraker and Robert K. Jaedicke. *Quantitative Analysis for Business Decisions*. Homewood, Illinois: Richard D. Irwin, Inc., 1961, 358 pp., $10.60. No. 6, p. 53.

Blum, Eleanor. *Sources for Reference in the Mass Media, An Annotated, Selective Bibliography*. Urbana, Illinois: Institute of Communications Research, University of Illinois, 1961, 46 pp., free but in short supply. No. 6, p. 53.

Brinton, James E. Deriving an Attitude Scale from Semantic Differential Data. *Public Opinion Quarterly*, Vol. 25, No. 2, Summer 1961, pp. 289-295. No. 6, p. 48.

Bristol, Lee (Editor). *Developing the Corporate Image*. New York: Charles Scribners Sons, 1960, 298 pp., $5.95. No. 1, p. 24.

Buckley, Earl A. *How to Increase Sales with Letters*. New York: McGraw-Hill Book Company, 1961, 182 pp. $5.00. No. 5, p. 38.

Burdick, Harry A., Edward J. Green and Joseph W. Loveland. Predicting Trademark Effectiveness. *Journal of Applied Psychology*, Vol. 43, No. 5, October 1959, pp. 285-286. No. 1, p. 28.

Couch, Arthur and Kenneth Keniston. Yeasayers and Naysayers: Agreeing Response Set as a Personality Variable. *Journal of Abnormal and Social Psychology*, Vol. 60, No. 2, March 1960, pp. 151-174. No. 4, p. 45.

Deming, W. Edwards. *Sample Design in Business Research*. New York: John Wiley & Sons, Inc., 1960, $12.00. No. 2, p. 26.

Donald, Marjorie N. Implications of Nonresponse for the Interpretation of Mail Data. *Public Opinion Quarterly*, Vol. 24, No. 1, Spring 1960, pp. 90-114. No. 1, p. 26.

Dunn, S. Watson. *Advertising: Its Role in Modern Marketing*. New York: Holt, Rinehart and Winston, 1961, 621 pp., $8.95. No. 6, p. 53.

Ehrenberg, A. S. C. The Pattern of Consumer Purchases. *Applied Statistics—A Journal of the Royal Statistical Society*, Vol. 8, No. 1, March 1959, pp. 26-41. No. 1, p. 26.

Ehrenberg, A. S. C. A Study of Some Potential Biases in the Operation of a Consumer Panel. *Applied Statistics—A Journal of the Royal Statistical Society*, Vol. 9, No. 1, March 1960, pp. 20-27. No. 1, p. 26.

Ehrlich, June S. and David Riesman. Age and Authority in the Interview. *Public Opinion Quarterly*. Vol. 25, No. 1, Spring 1961, pp. 39-56. No. 5, p. 36.

Evans, Franklin B. Psychological and Objective Factors in the Prediction of Brand Choice: Ford versus Chevrolet. *Journal of Business*. Vol. 32, No. 4, October 1959, pp. 340-369. No. 4, p. 44.

Fenton, John M. *In Your Opinion*. Boston: Little, Brown & Company, 1961, 220 pp., $3.95. No. 6, p. 53.

Fourt, Louis A. and Joseph W. Woodlock. Early Prediction of Market Success for New Grocery Products. *Journal of Marketing*, Vol. 25, No. 2, April 1960, pp. 31-38. No. 2, p. 32.

Goldfarb, Nathan. *An Introduction to Longitudinal Statistical Analysis*. Glencoe, Illinois: The Free Press, 1960. 220 pp. $5.00. No. 5, p. 38.

Graybill, Franklin A. *An Introduction to Linear Statistical Models, Volume I*. New York: McGraw-Hill Book Company, 1961. 463 pp. $12.50. No. 5, p. 38.

Gruen, Walter. Preference for New Products and Its Relationship to Different Measures of Conformity. *Journal of Applied Psychology*, Vol. 44, No. 6, December 1960, pp. 361-364. No. 3, p. 38.

Gulliksen, Harold and Samuel Messick (Editors). *Psychological Scaling*. New York: John Wiley and Sons, Inc., 1960. 211 pp. $5.00. No. 5, p. 38.

Hardy, Hugh. Study Shows Public Attitude Toward Advertising is Surprisingly Good. *Advertising Age*, Vol. 31, No. 49, December 5, 1960, pp. 77-82. No. 3, p. 36.

Himmelweit, Hilde T., A. N. Oppenheim and Pamela Vince. *Television and the Child*. London: Oxford University Press, 1958. No. 4, p. 41.

Holdren, Bob R. *The Structure of a Retail Market and the Market Behavior of Retail Units*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1960. 203 pp. $4.50. No. 5, p. 38.

Holmes, Parker M. *Marketing Research Principles and Readings*. Cincinnati, Ohio: Southwestern Publishing Co., 1960, 641 pp., $7.50. No. 6, p. 49.

Houseman, Earl E. and Benjamin Lipstein. Observation and Audit Techniques for Measuring Retail Sales. *Agricultural Economics*, Vol. 12, No. 3, July 1960, pp. 61-72. No. 2, p. 30.

Kemsley, W. F. F. Designing a Budget Survey. *Applied Statistics—A Journal of the Royal Statistical Society*, Vol. 8, No. 2, June 1959, pp. 114-123. No. 1, p. 28.

Kemsley, W. F. F. Interviewer Variability and a Budget Survey. *Applied Statistics*, Vol. 9, No. 2, June 1960, pp. 122-128. No. 4, p. 44.

Kent, A. J. and C. Nuttall. An Attempt to Answer the Question—Does Advertising Pay? *The Incorporated Statistician*, Vol. 10, No. 1, February 1960, pp. 29-35. No. 5, p. 35.

Kilpatrick, F. P. and Hadley Cantril. Self-Anchoring Scaling, A Measure of Individuals' Unique Reality Worlds. *Journal of Individual Psychology*, Vol. 16, No. 2, November 1960. No. 2, p. 28.

Kimball, Andrew E. Increasing the Rate of Returns in Mail Surveys. *Journal of Marketing*, Vol. 25, No. 6, October 1961, pp. 63-64. No. 6, p. 53.

Klapper, Joseph T. *The Effects of Mass Communication*. Glencoe, Illinois: The Free Press, 1960, 302 pp., $5.00. No. 3, p. 37.

Lee, A. M. and A. J. Burkart. Some Optimization Problems in Advertising Media Planning. *Operational Research Quarterly*, Vol. 11, No. 3, September 1960, pp. 113-122. No. 3, p. 34.

Lipstein, Benjamin. Tests for Test Marketing. *Harvard Business Review*, Vol. 39, No. 2, March/April 1961, pp. 74-77. No. 4, p. 43.

Lockley, Lawrence C. *Use of Motivation Research in Marketing*. New York: National Industrial Conference Board Inc., 1960, $1.50 to NICB associates, $7.50 to non-associates. No. 2, p. 27.

Lundstedt, Sven (Editor). *Consumer Behavior in 1961, A Summary Report*. Ann Arbor, Michigan: Foundation for Research on Human Behavior, 1961, 38 pp., $1.00. No. 5, p. 36.

Lyle, Jack. Semantic Differential Scales for Newspaper Research. *Journalism Quarterly*, Vol. 37, No. 4, Autumn 1960, pp. 559-646. No. 6, p. 48.

Madow, W. G., H. H. Hyman and R. J. Jessen. *Evaluation of Statistical Methods in Obtaining Broadcast Ratings: Report of the Committee on Interstate and Foreign Commerce*. Washington, D. C.: U. S. Government Printing Office, 1961. 163 pp. Free. No. 5, p. 38.

Maloney, John C. Portfolio Tests—Are They Here To Stay? *Journal of Marketing*, Vol. 25, No. 5, July 1961, pp. 32-37. No. 6, p. 53.

Miller, David W. and Martin Starr. *Executive Decisions and Operations Research*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1960, 446 pp. No. 1, p. 25.

Munn, Henry L. Brand Perception as Related to Age, Income, and Education. *Journal of Marketing*, Vol. 42, No. 3, January 1960, pp. 29-34. No. 1, p. 28.

Myers, Lawrence, Jr. and Eric F. Gardner. An Inexpensive Method to Determine the Efficiency of a Television Program. *Journal of Applied Psychology*, Vol. 44, No. 1, February 1960, pp. 39-42. No. 1, p. 25.

Nossiter, Bernard D. You've Never Been Quizzed by a Pollster? *The Washington Post*, October 16, 1960, p. E3. No. 2, p. 29.

Paranka, Stephen. Marketing Predictions from Consumer Attitudinal Data. *Journal of Marketing*, Vol. 25, No. 1, July 1960, pp. 46-51. No. 1, p. 27.

Parker, Edwin B. Subliminal Stimulation and Voting Behavior. *Journalism Quarterly*, Vol. 37, No. 4, Autumn 1960, pp. 588-590. No. 3, p. 35.

Paulu, Burton. *British Broadcasting in Transition*. Minneapolis: University of Minnesota Press, 1961. 250 pp. $5.00. No. 5, p. 38.

Quenouille, M. H. *Rapid Statistical Calculation*. London: Charles Griffin and Company, Ltd., 1959. 250 pp. $5.00. No. 5, p. 38.

Rokeach, Milton. *The Opened and Closed Mind*. New York: Basic Books, Inc., 1960, 448 pp., $7.50. No. 4, p. 46.

Sandage, C. H. (Editor). *The Promise of Advertising*. Homewood, Illinois: Richard D. Irwin, Inc., 1961, 207 pp., $5.50. No. 6, p. 53.

Sandage, C. H. and Vernon Fryburger (Editors). *The Role of Advertising: A Book of Readings*. Homewood, Ill.: Richard D. Irwin, Inc., 1960, 499 pp., $8.00. No. 2, p. 30.

Schettler, Clarence. *Public Opinion in American Society*. New York: Harper and Brothers, 1960. 534 pp. $7.00. No. 5, p. 38.

Schramm, Wilbur, Jack Lyle and Edwin B. Parker. *Television in the Lives of Our Children*. Stanford: Stanford University Press, 1961, 324 pp., $6.00. No. 6, p. 47.

Schwerin Research Corporation. When Should the Effects of Television Advertising Be Measured? Part I: Recall. *Technical and Analytical Review*, No. 4, Spring 1960, pp. 1-16. No. 2, p. 31.

Schwerin Research Corporation. When Should the Effects of Television Advertising be Measured? Part II: Changes in Attitude and Behavior. *Technical and Analytical Review*, No. 5, Summer 1960, pp. 1-14. No. 3, p. 36.

Scott, Christopher. Research on Mail Surveys. *Journal of the Royal Statistical Society, Series A*, Vol. 124, Part 2, 1961 pp. 143-205. No. 6, p. 50.

*Search*. Published bi-monthly by the National Project in Agricultural Communications. East Lansing: Michigan State University. $5.00 a year. No. 5, p. 34.

Selltiz, Claire, Marie Jahoda, Morton Deutsch, and Stuart W. Cook. *Research Methods in Social Relations* (Revised in one volume). New York: Henry Holt and Company, Inc., 1960. 622 pp., $5.50. No. 6, p. 50.

Slonim, Morris J. *Sampling in a Nutshell*. New York: Simon and Schuster, 1960, 144 pp., $3.50. No. 6, p. 53.

Summers, Harrison B. Qualitative Information Concerning Audiences of Network Television Programs. *Journal of Broadcasting*, Vol. 5, No. 2, Spring 1961, pp. 147-160. No. 5, p. 37.

Thompson, J. Walter, Company. *Population and Its Distribution. The United States Markets* (Eighth Edition). New York: McGraw-Hill, 1961, 471 pp., $24.00. No. 6, p. 53.

Tucker, W. T. and John J. Painter. Personality and Product Use. *Journal of Applied Psychology*, Vol. 45, No. 5, October 1961, pp. 325-329. No. 6, p. 53.

United States Information Agency—Office of Research and Analysis. *Post-Summit Trends in British and French Opinion of the United States and the Soviet Union, June 1960*. In the New York *Times*, October 27, 1960. No. 2, p. 33.

Weinberg, Robert S. *An Analytical Approach to Advertising Expenditure Strategy*. New York: Association of National Advertisers Inc., 1960, 125 pp., $2.50 to ANA members, $5.00 to non-members. No. 1, p. 23.

Weinberger, Martin. Does the "Sleeper Effect" Apply to Advertising? *Journal of Marketing*, Vol. 25, No. 6, October 1961, pp. 65-67. No. 6, p. 53.

Winick, Charles. The Relationship Among Personality Needs, Objective Factors, and Brand Choice: A Re-examination. *Journal of Business*, Vol. 34, No. 1, January 1961, pp. 61-66. No. 4, p. 44.

Wood, James P. *Advertising and the Soul's Belly, Repetition and Memory in Advertising*. Athens: University of Georgia Press, 1961, 116 pp., $3.50. No. 4, p. 42.

Zacher, Robert V. *Advertising Techniques and Management*. Homewood, Illinois: Richard D. Irwin, Inc., 1961, 618 pp., $10.00. No. 6, p. 53.

## FEDERAL STATISTICS IN ADVERTISING

Ingrid C. Kildegaard, ARF Research Statistician


| | |
|---|---|
| The Consumer Price Index | No. 1, p. 22 |
| Small Area Statistics | No. 2, p. 25 |
| Consumer Possessions and Buying Intentions | No. 3, p. 27 |
| The 1958 Census of Business | No. 4, pp. 34-35 |
| County Business Patterns | No. 5, pp. 26-27 |
| Consumer Income Data | No. 6, pp. 37-39 |


## ON METHODS

Gwyn Collins, ARF Research Mathematician


| | |
|---|---|
| Attitude Scaling | No. 3, pp. 28-33 |
| Correlation and Regression | No. 4, pp. 36-40 |
| Factor Analysis | No. 5, pp. 28-32 |
| Analysis of Variance | No. 6, pp. 40-46 |


## EDITORIALS

Charles K. Ramond, ARF Technical Director


| | |
|---|---|
| Science in Wonderland | No. 1, p. 32 |
| Uniquack on Advertising Research | No. 2, p. 36 |
| Agostini's Constant | No. 3, p. 40 |
| The Girls Who Can't Say No | No. 4, p. 48 |
| Keep It Crisp | No. 5, p. 40 |
| The Balance Sheet | No. 6, p. 56 |

# ARF and Its Objectives

The Advertising Research Foundation is a non-profit organization supported by advertisers, advertising agencies and advertising media. Its basic purpose is to further scientific practices and promote greater effectiveness in advertising and marketing through objective and impartial research.

Specifically, the Foundation offers consultative and professional guidance with regard to the conduct of research studies; supervises the technical aspects of advertising research projects; and appraises published research studies conducted by others. Further, the ARF seeks to develop new research methods and techniques; to analyze existing techniques and define their proper applications and limits of usefulness; and to establish research standards and criteria.

The ARF collects and disseminates advertising and marketing data for the benefit of members. It is interested in exploring any research plan which might increase the industry's knowledge of how to make advertising more effective.

# Officials and Directors of ARF

## OFFICIALS

ARTHUR H. HAYES, Chairman*
President, CBS Radio, A Division of Columbia Broadcasting System, Inc.

FRANK W. MANSFIELD, Vice Chairman*
Director of Marketing Research, Sylvania Electric Products Inc.

LYNDON O. BROWN, Treasurer*
Senior Vice President, Dancer-Fitzgerald-Sample, Inc.

* Also a Director.

ALCUIN W. LEHMAN, President*
Advertising Research Foundation, Inc.

CHARLES K. RAMOND, Technical Director
Advertising Research Foundation, Inc.

CYRUS C. YOUNG, Director of Development and Services
Advertising Research Foundation, Inc.

GILBERT H. WEIL, General Counsel

## DIRECTORS

PETER W. ALLPORT
President, Association of National Advertisers, Inc.

RICHARD J. BABCOCK
President, Farm Journal, Inc.

HUGH M. BEVILLE, JR.
Vice President—Planning & Research
National Broadcasting Company

ARTHUR C. BRYAN
President, Union Carbide Consumer Products Company

REX M. BUDD
Director of Advertising, Campbell Soup Company

THOMAS C. DILLON
Executive Vice President & Treasurer
Batten, Barton, Durstine & Osborn, Inc.

WILLIAM R. FARRELL
Director of Marketing Services
Monsanto Chemical Company

FREDERIC R. GAMBLE
President, American Association of Advertising Agencies, Inc.

HAROLD J. GRAHAM
Assistant Secretary, Hartford Fire Insurance Company Group

JOHN W. HARTMAN
President, Bill Brothers Publications

DONALD M. HOBART
Senior Vice President and Director of Research
The Curtis Publishing Company

ARNO H. JOHNSON
Vice President, J. Walter Thompson Company

ROBERT E. KAHL
Director of Marketing Services, The Borden Company

WALTER C. KURZ
Vice President & Director of Sales, Chicago Tribune

PETER LANGHOFF
Vice President & Director of Research
Young & Rubicam, Inc.

A. EDWARD MILLER
Publisher, McCall's

CHARLES A. POOLER
Senior Vice President, Benton & Bowles, Inc.

CHARLES L. RUMRILL
President, The Rumrill Company, Inc.

HARRY F. SCHROETER
Director, General Advertising Department
National Biscuit Company

JAMES M. WALLACE
Vice President, N. W. Ayer & Son, Inc.

WALLACE H. WULFECK
Vice Chairman, William Esty Company, Inc.

3473